A TETRAZINE-TERMINAL ALKENE SYSTEM FOR PROFILING PROTEIN FATTY ACYLATION MODIFICATIONS

A Thesis

by

SASHA MALA DEONARINE-CHIHAK

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Wenshe Liu
Committee Members,	Frank Raushel
	Marcetta Darensbourg
	Ann Kier
Head of Department,	Simon North

December 2016

Major Subject: Chemistry

ABSTRACT

Protein fatty acylation is included in the class of modifications where attachment via an acyl linkage is employed including myristoylation and palmitoylation. This work focuses on applying the novel tetrazine moiety to covalently target alkene functionalized fatty acids via a highly sensitive and selective inverse electron demand Diels alder cyclo-addition (iEDDA) reaction with accompanied visualization via in gel fluorescence imaging or western blot analysis.

This strategy was used in this study to compare the incorporation of the terminal alkene versus alkyne fatty acids mimics, by labeling of the modified proteins via iEDDA and Copper assisted Azide-Alkyne Cycloadditon (CuAAC), respectively. The successful labeling of the incorporated terminal alkene fatty acid provides another strategy for investigating fatty acyl modifications and so assist in validating the established library of curated fatty acyl modifications. Additionally, while differences in the location and intensity of labeling was observed between the alkene and alkyne fatty acid surrogates utilized in this study it was impossible to formulate conclusions on the effectiveness of the iEDDA labeling method over the traditionally used CuAAC strategies without quantitative mass spectroscopy data.

DEDICATION

For my husband and our family

ACKNOWLEDGEMENTS

To my parents, Winston and Usha Deonarine, thank you for starting me along my journey through science, and to Chuck and Regina Chihak, for your constant support along the road. As reach another fork in the road, I would like to thank my undergrad research advisor, Dr. N. Carmosini, and the UWL's McNair scholars program, especially Jessica Thill for their assistance and encouragement in pursuing a graduate education.

I would like to thank my committee chair, Dr. Wenshe Liu, and committee members, Dr. Frank Raushel, Dr. Marcetta Darensbourg and Dr. Ann Kier for their guidance in completing my research and acknowledge the support of the Welch foundation (A1715) for providing the necessary funding.

My sincerest gratitude to the members of the organic chemistry laboratory program; Carrie Nicols, Janet Robinson, Dr. K. Harding, Dr. T. Tiner, Dr. A. Schaefer, as well as Sandy Horton and Dr. J Pennington. I am blessed to have had your guidance and encouragement through discovering my love of teaching in addition to your emotional support through the challenges.

Thank you to my friends; John Bescup, Dr. Betsy Brown and Katie Challis, your encouraging words were life savers when I was drowning in doubt. Lastly, to my husband, Vince Chihak, for your unwavering dedication to our relationship, your honest feedback and those fluffy pancakes (+100 points to Gryffindor). Thank you.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

INTRODUCTION

Post translational modifications (PTM) include the array of chemical and enzymatic alterations to amino acid residues which augment the chemical diversity of proteins[1]. These alterations can range from the attachment of simple functional groups to the addition of small molecules and macromolecules. Protein fatty-acylation describes one such modification; it involves the covalent attachment of assorted fatty acids onto the amino acid side chain residues of cysteine, serine, threonine, N^{ε}-lysine and N-terminal amines. These transformations have been shown to affect the subcellular localization and activity of protein-protein and protein-membrane interactions in a wide variety of biological processes and associated disease states[2].

In one example, the palmitoylation (16-carbon fatty acid modification) of Wnt proteins was identified as essential for its proper functioning as it increased the ….with defective Wnt signaling implicated in several forms of cancer and degenerative diseases[3]. Similarly, myristoylation (14-carbon fatty acid modification) of structural proteins on HIV-1 (Human Immunodeficiency Virus 1) has been shown to affect its viral infectivity [4]. These examples illustrate the key role that protein fatty acylation play in cellular homeostasis and underscore the importance of determining their amenability to serve as either disease biomarkers, for diagnostic applications, or as therapeutic targets for pharmacological applications.

Dietary fats are obtained from a variety of sources and consist mostly of long chain fatty acids. These are highly condensed energy sources, contributing approximately 30-40% of an individual's daily energy intake. Those with chain lengths more than three carbons are enzymatically converted into fatty acyl-CoAs by their corresponding acyl-CoA synthetases [5]. These undergo repeated cycles of β-oxidation to produce acetyl-CoA that enters into the tricarboxylic acid cycle (TCA), eventually yielding intracellular energy carrier molecules such as adenosine triphosphate (ATP). The variety of acyl-Co-A intermediates formed over the course of β-oxidation are high-energy labile thioesters well suited to participate in nucleophilic acyl substitution reactions with protein nucleophiles.(Figure 1)

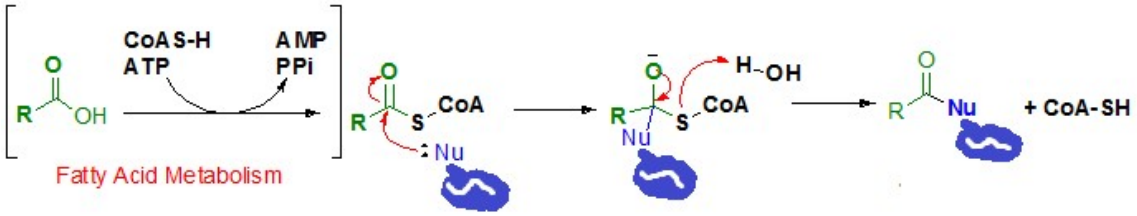


Figure 1: A diagrammatic summary of the mechanism of fatty acylation. Nu: represents the amino acid side chain functional groups of serine, cysteine and Nε-lysine as well as glycine free amine. R represents the variety of fatty acid chain lengths observed on acyl-Co-A intermediates formed via the β-oxidation pathway.

The challenges of studying fatty acid modifications has been well documented[6],[7]. The lack of specific antibodies for fatty acid modifications has resulted in initial investigations being limited to the use of radioactively labeled chemical reporters, typically [H^3] or [I^{125}] containing fatty acids[8]. These often require lengthy

exposure times (1-3 months), produce weak signals and in certain cases require high financial investment [9]. Furthermore, the difficulties associated with chromatographic separation of highly hydrophobic fatty acylated proteins, the poor ionization of samples and low sensitivity of mass spectroscopy detection have made these traditional research methods almost impractical [7].

This has prompted a shift towards the use of small molecule techniques. Small molecule strategies involve the incorporation of functionalized chemical reporter molecules followed by reaction with a complementary capture probe to facilitate the visualization or isolation of the pre-targeted cellular components from the complex biological milieu [9],[10], [11] . Developments in bio-orthogonal reaction chemistry has provided a range of different functionalities as potential fatty acid mimics. For instance, chemical reporters with chloroacetyl, terminal alkyne, terminal azide and strained alkenes were identified as possible functionalities to serve as selective reporter molecules. It was found that the application of a chloroacetyl reporter resulted in high cellular toxicity and so could not be used in direct cellular application. The strained alkenes on the other hand showed low incorporation and so were not suitable for biological applications [12],[13].

The use of terminal alkyne or azide functionalized fatty acids however has proved successful. This example has found the widest application and greatest documented success [14]. Once incorporated into the cellular proteome the installed azide or alkyne is coupled via the copper assisted azide alkyne cycloaddition (CuAAC) with a complementary capture probe. Or in addition terminal azides may also be targeted using

phosphines via the Staudinger ligation reaction. The aqueous compatibility and rapid reaction kinetics of the azide-alkyne labeling strategy has encouraged their utility in the profiling of fatty acylation despite their noted spatial and structural difference from native dietary fatty acids[14]. (Figure 2) Note that the azide and alkyne surrogates show a clear spatial and structural difference from the native fatty acid particularly in their terminal functionalty. The azide and alkyne are significantly longer and less hydrophobic than the terminal alkyl group of the myristic acid and shows better similarity to the 13-tetradecenoic acid in this respect.

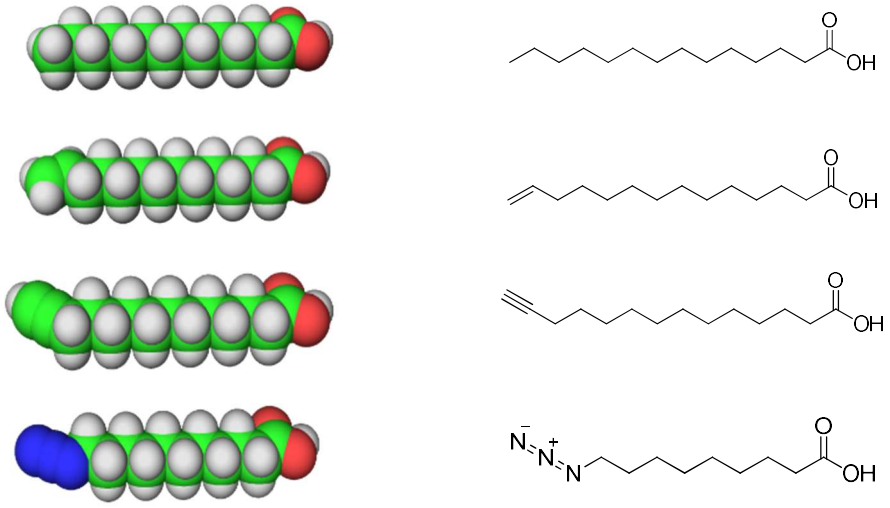


Figure 2: A space-filled and line representations of myristic acid and its alkene, alkyne and azide surrogates - 13-tetradecenoic acid, 13-tetradecynoic acid and 12-azidododecanoic acid. [Green – carbon, white- hydrogen, red- oxygen, blue- nitrogen]

While a large dataset of proteomic profiling studies on fatty acylation (palmitoylation) has been curated at SwisPalm (http://swispalm.epfl.ch/), only a meager 10% has been validated[7]. In addition a number of false positive results have identified within this dataset indicating a need for the development of alternative strategies for not

only profiling fatty acyl modifications but also validating the results which has been accumulated. Thus, while azides and alkynes may appear to serve as mimics of fatty acid substrates their structural differences may in reality result in a non-canonical metabolic process leading to partial or improper fatty acylated targets [15]. This necessitates the development of alternative strategies to not only validate current data sets of fatty acylated modifications but also to offer improved methodologies for profiling of protein post translational fatty acid modifications. The similarity of a terminal alkene fatty acid mimic both in shape and hydrophobicity to the corresponding native fatty acids facilitates their incorporation into the innate fatty acid metabolic pathways and reduces the likelihood of off target acylation.

Furthermore, the use of a terminal olefin reporter molecule facilitates the application of a tetrazine capture probe, which provides a selective bio-orthogonal conjugation mechanism. Additionally, the probe may be functionalized for rapid visualization using a fluorescent moiety or a biotin tag for enrichment and isolation of the modified proteome[16].

A CHEMICAL-BIOLOGY STRATEGY: TETRAZINE + TERMINAL ALKENE

THE FATTY ACID MIMIC

The similarity in shape and hydrophobicity of the terminal-olefin chemical reporter to corresponding native fatty acids suggests that they would serve as successful and efficient mimics of native fatty acids. This should allow these small molecules to seamlessly be accepted by endogenous enzymes and so provide a better representation of fatty acylation activities.

This study utilized only even numbered long-chain terminal-olefin fatty acids as odd numbered acids were not expected to lead to significantly different labeling results. The fatty acids, which were used were obtained from two sources; those with 6-10 carbons were obtained via commercial sources, while those with C12, C14, and C16 were synthesized in the laboratory **(Figure 3).** In order to determine the effectiveness of our terminal alkene reporter a similar C14 terminal alkyne fatty acid reporter, was prepared **(Figure 4).**

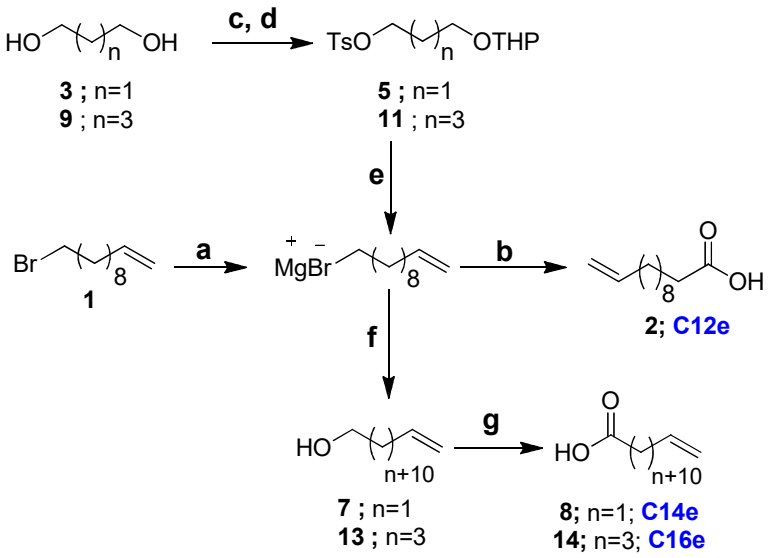


Figure 3: The synthetic strategy to prepare the terminal-olefin fatty acids C12e, C14e, C16e for metabolic labeling. Using the commercially available 11-bromo-1-undecene, via the following steps: a. Magnesium, Iodide, anh. THF; b. CO_2; c. DCM, DHP, PTSA, 0oC - r.t.; d. anh. DCM, anh. Pyridine, pTSCl, r.t.; e. anh. THF; f. Li_2CuCl_4, r.t.; g. Acetic acid, CrO3, H2O, 0^{o}C-r.t.

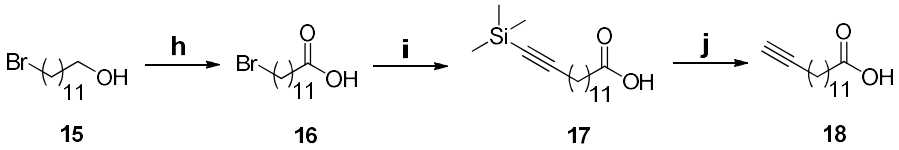


Figure 4: The synthetic strategy to prepare the terminal- alkyne fatty acid C14y for metabolic labeling. Starting with the commercially available 12-bromo-1-dodecanol according to the following steps: h. Acetic acid, CrO_3, H_2O, 0^{o}C-r.t.; i. TMS-acetylene, n-BuLi, anh THF, -78 ^{o}C-r.t. j. K_2CO_3, MeOH, r.t.

THE CHEMICAL REPORTER

In selecting the tetrazine chemical reporter for this study we reflected on the results published by Lee et. al. where a benzoic acid substituted tetrazine (**35a**) reacted successfully and with good reaction kinetics to a number of unstrained terminal alkenes.

In addition, Lee et. al also showed that a biaryltetrazine **(35b)** was effective in avoiding non-specific labeling of free cysteine side chains on BSA[16]. Furthermore, interest in the biaryltetrazine was shown by Pipkorn and coworkers who demonstrated that the biaryltetrazine (**35b**) proved compatible with aqueous solutions containing amines and disulfides [17]. Meanwhile Karver et. al. determined that the aqueous solubility of the tetrazine was 0.2 mM as a result stock solutions were prepared in DMSO and kinetic evaluation of the iEDDA reaction used a maximum concentration of only 5 μM of the tetrazine dye[18].

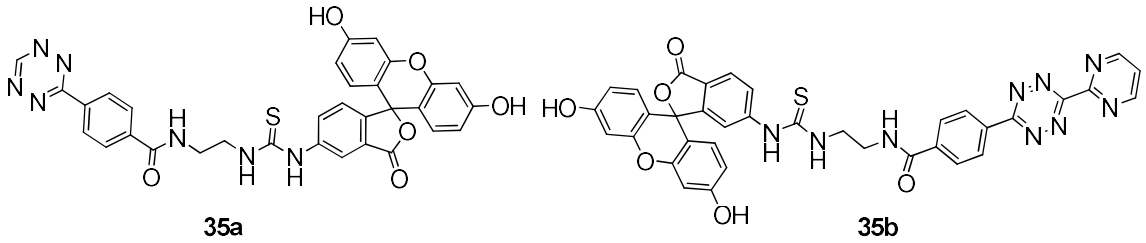


Figure 5: The chemical structure of the tetrazine molecules studied by Lee et. al.[16]

As the reaction kinetics for the biaryltetrazine containing a pyrimidyl and benzoic acid substituent has not yet been published, tetrazine **33** (**Tz-F**) was prepared, to expedite this process. Furthermore, to achieve the objective of protein enrichment a biotin conjugated analog, **32** (**Tz-B**) was also prepared (**Figure 6)**.

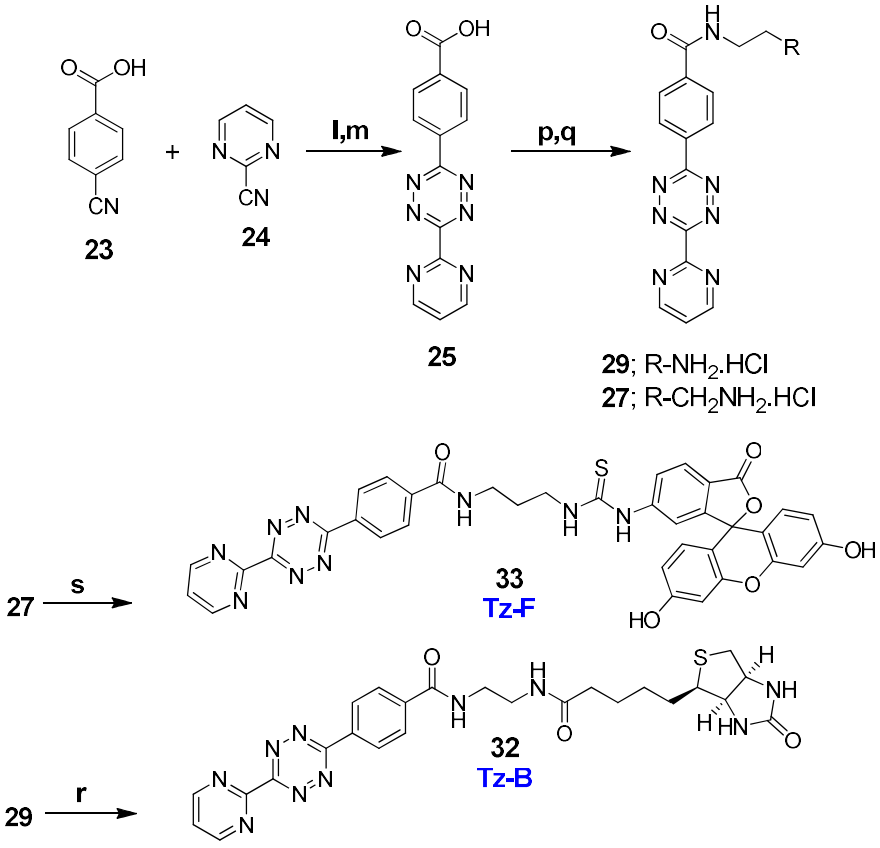


Figure 6: The synthetic scheme for the preparation of chemical reporters 32 [Tz-B] and 33 [Tz-F] used for targeted labeling of fatty acyl modifications. Reagents and conditions: l. $NH_2NH_2.H_2O$, anh. EtOH, reflux 90°C; m. Acetic acid, $NaNO_2$, H_2O, 10 °C; p. anh DMF, HATU, DIPEA, r.t. q. anh MeOH, 4M HCl in 1,4 dioxane, 0 °C-r.t.; r. anh DMF, DIPEA, NHS-Biotin , r.t.; s. anh MeOH/THF 2:1, TEA, FITC, 0 °C-r.t.

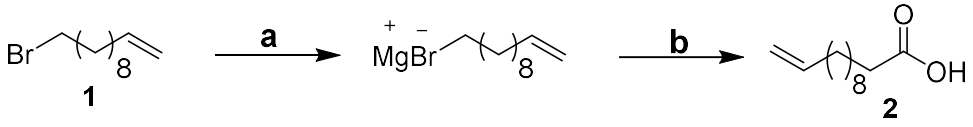
MATERIALS AND METHODS

General information

All reactions involving moisture sensitive reagents were performed using oven dried glassware under nitrogen atmosphere. Anhydrous solvents were obtained from

commercial sources or via distillation using standard laboratory protocols. Analytical thin-layer chromatography (TLC) was performed on EMD Millipore silica gel 60 F254 plates. Visualization was accomplished by UV irradiation at 254 nm or by staining with ninhydrin (0.3% w/v in glacial acetic acid/n-butyl alcohol 3:97). Flash column chromatography was performed with flash silica gel (particle size 32-63 μm) from Dynamic Adsorbents Inc. (Atlanta, GA). Proton NMR spectra were obtained on Varian 300 and 500 MHz NMR spectrometers. Chemical shifts are reported as δ values in parts per million (ppm) as referenced to the residual solvents: deuterated chloroform (7.27 ppm for ^{1}H and 77.23 ppm for ^{13}C); deuterium oxide (4.80 ppm for ^{1}H); dimethyl sulfoxide (2.50 ppm for ^{1}H and 39.51 for ^{13}C). ^{1}H NMR spectra are tabulated as follows: chemical shift, (multiplicity; s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), number of protons).

Synthesis of 12- carbon terminal alkenyl fatty acid [C12e]

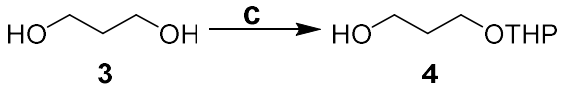


(1)

The dodec-11-enoic acid was prepared in a similar fashion as reported by Wube et al.[19]. The halide **1** (1 eq, 0.5 mL, 2.0 mmol) was added to a stirring mixture of activated magnesium (2.5 eq, 138 mg, 5.1 mmols) and iodide (catalytic) in anhydrous tetrahydrofuran under inert atmosphere while gently heating. When the solution was

decolorized heating was stopped and carbon dioxide was bubbled through the reaction mixture to give a thick creamy mixture. The reaction was quenched by the addition of chilled saturated ammonium chloride solution (10 mL) and diluted with diethyl ether (20 mL). The organic phase was separated and sequentially washed with water (20 mL), saturated sodium bicarbonate (20 mL) and saturated sodium chloride (20 mL), dried with sodium sulfate and concentrated via rotary evaporation. Purification was achieved via silica gel column chromatography eluting with a mixture of hexanes/ ethyl acetate to obtain **3** (300 mg, 67%) as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.85 (m, 1H), 5.05-4.99 (m, 2H), 2.37 (m, 2H), 2.08 (m, 2H), 1.68 (m, 2H), 1.31 (m, 14H).

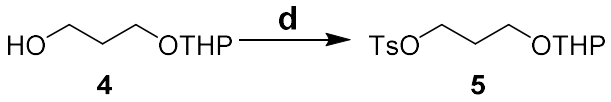
Synthesis of 14- carbon terminal alkenyl fatty acid [C14e]



3 **4** (2)

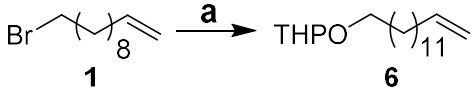
Propane diol was protected with DHP using the strategy described by Guillod et. al. [20].To a solution of propane diol, **3** (2 eq, 5.34 mL, 78.8 mmols) in anhydrous DCM (20 ml) was added p-toluene sulfonic acid (0.4 eq, 300 mg, 1.7 mmols) portion-wise. The solution was cooled to 0 ºC and DHP (1 eq, 3.8 mL, and 41.7 mmols) was added dropwise over 2 hours and the mixture warmed to room temperature. The reaction mixture was washed with water, dried with anhydrous $NaSO_4$ and purified via silica column chromatography eluting with 70:30 hexanes/ethyl acetate. The product, **4** was

collected as a colorless oil (2.88 g, 90%) ^{1}H NMR (300Mz, $CDCl_3$) δ 4.61 (s, 1H), 3.85 (m, 2H), 3.53 (m, 2H), 1.92 -1.54 (m, 8H).



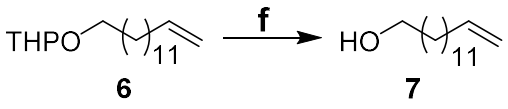
(3)

The mono protected alcohol was further protected using the method published by Guillod et. al. [20]. To the solution of **4,** (1 eq, 1.67 g, 10.4 mmols) in anhydrous DCM (10 mL) under inert atmosphere conditions was added anhydrous pyridine (2 eq, 1.68 mL, 20.8 mmols). Toluene sulfonyl chloride (1.25 eq, 2.49 g 15.6 mmols) was dissolved in anhydrous DCM (5 mL) and added portion wise at room temperature to give a yellow solution. Consumption of alcohol **5** was confirmed via TLC. The reaction mixture was then diluted with a portion of DCM (25 mL), sequentially washed with 5% HCl solution (2 x 30 mL), saturated $NaHCO_3$ solution (2 x 30 mL), saturated NaCl solution (2 x 30 mL), dried with $MgSO_4$, filtered and concentrated via rotary evaporation. Purification was performed via silica gel column chromatography eluting with 70:30 hexanes/ethyl acetate to obtain **5** (1.6 g, 49%), as yellow oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 7.81 (d, 2H), 7.36 (d, 2H), 4.45(s, 1H), 4.14 (m, 4H), 3.81 (m, 2H), 3.45 (m, 2H), 2.48 (s, 3H), 1.95 (m, 2H), 1.85-1.48 (m, 8H).



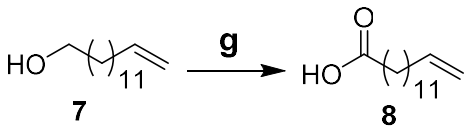
1 **6** (4)

The grignard reaction of the protected diol with the alkenyl halide was performed as documented by Rezanka and coworkers [21]. A solution of halide **1** (1.5 eq, 0.3 mL, 1.3 mmol) in anhydrous THF (5 mL), was added to a stirring mixture of activated magnesium (2.5 eq, 56 mg, 2.3 mmols) and iodide (catalytic) in anhydrous THF(10 mL) under inert atmosphere while gently heating. When the solution was decolorized heating was stopped and the protected diol **5** (1 eq, 290 mg, 0.9 mmols) was added dropwise. The solution was stirred for 3 hours unheated following which Li_2CuCl_4 (0.2 eq, 0.2 mL, 0.18 mmols) was added. The resulting green mixture continued to stir at room temperature overnight to yield a yellow brown mixture. This was sequentially extracted with chilled $NH_4^+Cl^-$ (20 mL), water (40 mL), saturated $NaHCO_3$ (20 mL), saturated NaCl (20 mL), dried with Na_2SO_4 and concentrated via rotary evaporation to give a mixture of **6** and **7**. Purification via silica gel column chromatography elution with 70:30 hexanes/ethyl acetate isolated **7** (80 mg, 40%) and **6** (196 mg, 72%) as colorless oils. ^1H NMR (300Mz, $CDCl_3$) δ 5.85 (m, 1H), 5.05-4.94 (m, 2H), 4.61 (m, 1H), 3.90 (m, 2H), 3.42 (m, 2H), 2.05 (m, 2H), 1.58 (m, 6H), 1.42-1.32 (m, 16H).



6 **7** (5)

p- Toluene sulfonic acid (0.05 eq, 60 mg, 0.34 mmols) was added to **6** (1 eq, 196 mg, 6.6 mmols) dissolved in methanol (10 mL) and stirred overnight at room temperature. Sodium hydroxide (1M, 0.5mL) was added and the reaction was concentrated. The residue was re-dissolved in ethyl acetate and sequentially washed with water (10 mL), saturated $NaHCO_3$ solution (10 mL), saturated NaCl solution (10 mL), dried with sodium sulfate and concentrated. The crude oil was purified via silica gel column chromatography 90:10 hexane/ ethyl acetate to give **7** (115 mg, 85%), as a colorless oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 5.85 (m, 1H), 5.05-4.94 (m, 2H), 3.67 (t, 2H), 2.37 (t, 1H), 2.05 (m, 2H), 1.59 (m, 4H), 1.42-1.32 (m, 16H).

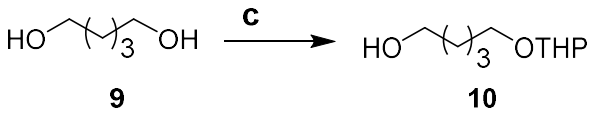


7 **8** (6)

Acetic acid (2.1 mL) was added to an ice cold solution of CrO_3 (3.9 eq, 201 mg, 3.1 mmols) in water (1 mL) to prepare the Jones' reagent. The alcohol **7** (1 eq, 110 mg, 0.52 mmols) was dissolved in acetone (0.5 mL) and cooled to 0^{o}C before being added drop wise to the Jones' solution maintained at 0^{o}C. Following complete addition of the alcohol the mixture was allowed to warm to room temperature and continuously stirred overnight. The dark solution was diluted with water and the pH was adjusted to ~8-10.

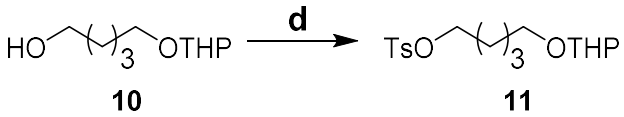
This was extracted with diethyl ether (20 mL) and the aqueous phase was separated. The aqueous phase was pH adjusted to ~4-6 and extracted with ethyl acetate (20 mL). The ethyl acetate phase was dried with Na_2SO_4 and concentrated to yield the alkenyl acid **8** (99.5 mg, 85%), as a colorless oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 5.85 (m, 1H), 5.05-4.94 (m, 2H), 3.67 (m, 2H), 2.05 (m, 2H), 1.65 (m, 2H), 1.37 (m, 16H).

Synthesis of 16- carbon terminal alkenyl fatty acid [C16e]



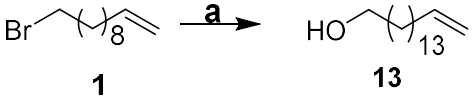
9 **10** (7)

Pentane diol was protected with DHP using the strategy described earlier [20].To solution of **9** (1.5 eq, 6 mL, 57.6 mmols) in anhydrous dichloromethane (20 ml) was added p-toluene sulfonic acid (0.08 eq, 550 mg, 3.2 mmols) portion-wise. The solution was cooled to 0 ^{o}C and dihydropyran (1 eq, 2.7 mL, and 38.4 mmols) was added dropwise over 2 hours as the mixture warmed to room temperature. The reaction mixture was washed with water, dried with anhydrous sodium sulfate and purified via silica column chromatography eluting with 70:30 hexanes/ ethyl acetate. The product, **10** was collected as a colorless oil (5.3 g, 73%) ^{1}H NMR (300Mz, $CDCl_3$) δ 4.59 (s, 1H), 3.85-3.36 (m, 6H), 1.86-1.41 (m, 14H).



10 **11** (8)

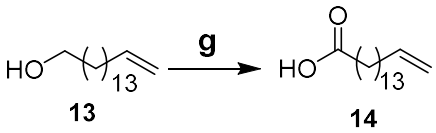
A solution of **10,** (1 eq, 5.3 g, 28.2 mmols) in anhydrous DCM (10 mL) under inert atmosphere conditions was chilled and triethylamine (3 eq, 12.2 mL, 84.6 mmols) was added dropwise. The solution was stirred for 10 min and a chilled solution of p-toluene-sulfonyl chloride (1 eq, 5.4 g, 28.2 mmols) in anhydrous DCM (5 mL) was added portion wise. The reaction was allowed to continue overnight warming to room temperature. A portion of DCM (25 mL) was used to dilute the reaction mixture and the reaction was quenched with 5% HCl solution (50 mL). The organic phase was washed with saturated $NaHCO_3$ solution (30 mL), saturated NaCl solution (30 mL), dried with $MgSO_4$, filtered and concentrated via rotary evaporation. Purification was performed via silica gel column chromatography eluting with 80:20 hexanes/ ethyl acetate to obtain **11** (6.0 g, 62%) as a colorless oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 7.80-7.33 (d, 2H), 7.36 (d, 2H), 4.57-4.54 (m, 1H), 4.55-4.03 (m, 1H) 3.80-3.73 (m, 2H), 3.53-3.37 (m, 2H), 2.47 (s, 3H), 1.86-1.52 (m, 14H).



1 **13** (9)

The grignard reaction of the protected pentane diol with the alkenyl halide was performed as detailed earlier [21]. A solution of halide **1** (1 eq, 2 mL, 9.1 mmol) in

anhydrous THF (5 mL), was added to a stirring mixture of activated magnesium (5 eq, 1.1 g, 45.3 mmols) and iodide (catalytic) in anhydrous THF(10 mL) under inert atmosphere while gently heating to reflux. When the solution was decolorized heating was continued for 3 hrs following which the mixture was allowed to cool slightly before the protected diol **11** (1 eq, 3 g, 9.5 mmols) dissolved in THF was added dropwise. Once the addition was completed the reaction mixture was stirred for another hour and Li_2CuCl_4 (0.2 eq, 0.2 mL, 0.18 mmols) was added. The resulting greenish mixture was allowed to stir at room temperature overnight to yield a yellowish mixture. This was washed with chilled $NH_4^+Cl^-$ (20 mL), water, saturated $NaHCO_3$, saturated NaCl, dried with Na_2SO_4 and concentrated via rotary evaporation to give **13**. Purification was completed via silica gel column chromatography elution with 90:10 hexanes/ ethyl acetate to yield **13** (440 mg, 15%) as a colorless oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 5.85 (m, 1H), 4.97 (m, 2H), 3.67 (m, 2H), 2.07 (m, 2H), 1.57 (m, 4H), 1.31 (m, 18H).



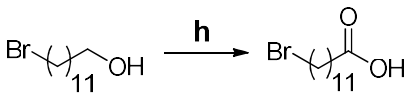
13 **14** (10)

Acetic acid (3.1 mL) was added to an ice cold solution of CrO_3 (3.9 eq, 308 mg, 3.1 mmols) in water (1 mL) to prepare the Jones' reagent. The alcohol **13** (1 eq, 190 mg, 0.80 mmols) was dissolved in acetone (0.8 mL) and cooled to 0°C before being added dropwise to the Jones' solution. The reaction mixture was allowed to warm to room temperature, while continuously stirring overnight. The dark solution was diluted with

water and the pH was adjusted to ~8-10. This was extracted with diethyl ether (30 mL) and the aqueous phase was separated. The aqueous phase was pH adjusted to ~4-6 and extracted with ethyl acetate (30 mL). The organic layer was dried with Na_2SO_4 and concentrated to yield the alkenyl acid **14** (180 mg, 89%), as a colorless oil.

^{1}H NMR (300Mz, $CDCl_3$) δ 5.73 (m, 1H), 4.87 (m, 2H), 2.28 (m, 2H), 1.95 (m, 2H), 1.56 (m, 2H), 1.31 (m, 20H).

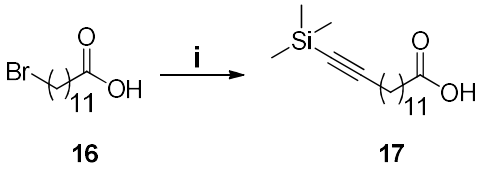
Synthesis of 14 carbon terminal alkynl fatty acid [C14y]



15 **16** (11)

The bromo acid was prepared as described by Wube et. al, acetic acid (13.1 mL) was added to an ice cold solution of CrO_3 (3.9 eq, 1.47 g, 18.5 mmols) in water (3 mL) to prepare the Jones' reagent, a yellow-orange solution[19]. The alcohol **15** (1 eq, 1.0 g, 3.8 mmols) was dissolved in acetone (3.75 mL) and cooled to 0ºC before being added dropwise to the Jones' solution. The reaction mixture was maintained at 0ºC for 2 hours before being allowed to warm to room temperature, while continuously stirring. The dark solution was diluted with water and the pH was adjusted to ~8-10. This was extracted with diethyl ether (20 mL) and the aqueous phase was separated. The aqueous phase was pH adjusted to ~4-6 and extracted with ethyl acetate (30 mL). The organic

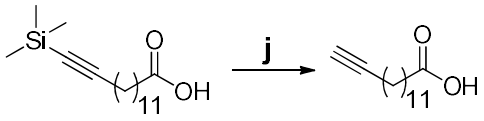
layer was dried with Na_2SO_4 and concentrated to yield the alkenyl acid **14** (806 mg, 82%), as a colorless oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 10.6 (br, 1H), 4.15 (m, 2H), 3.44 (m, 2H), 2.38 (m, 2H), 2.05 (m, 2H), 1.84 (m, 2H), 1.64 (m, 2H), 1.43 (m, 2H), 1.29 (m, 8H).



16 **17** (12)

To a solution of TMS acetylene (2.5 eq, 0.64 mL, 4.5 mmols) in anhydrous THF (5mL) at -78ºC under inert atmosphere was slowly added n-butyl lithium [1.6M/ Hexanes] (3 eq, 3.36 mL, 5.37 mmols)[19, 22]. The reaction mixture was removed from the dry ice bath and allowed to stir for 2 mins as a solution of **16** (1 eq, 500 mg, 1.79 mmol) in anhydrous THF (2.5 mL) was prepared. The reaction mixture was returned to the dry ice bath and the solution of **16** was added followed by HMPA (20 eq, 6.2 mL, 35.8 mmols) dropwise. The creamy beige mixture was stirred at -78ºC warming to room temperature to produce first a bronze-brown mixture continuing to darken to a dark brown almost black appearance at room temperature. This was allowed to stir overnight. The reaction mixture was again cooled to -78ºC and quenched by the addition of chilled saturated NH_4Cl solution. This mixture was extracted with DCM (3 x 20 mL) and the combined organic layers was washed with saturated NaCl solution (3 x 20 mL), dried with $MgSO_4$ and concentrated by rotary evaporation. The crude product was purified via

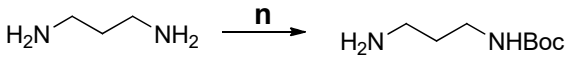
silica gel column chromatography eluting with 70:30 hexanes/ ethyl acetate to yield **17** (112 mg, 21%). ^{1}H NMR (300Mz, $CDCl_3$) δ 2.17 (m, 2H), 2.06 (m 2H), 1.96 (m, 4H), 1.36 (m, 4H), 1.12 (m, 10H), 0.01 (m, 9H).



17 **18** (13)

To a solution of **17** (1 eq, 112 mg, 0.38 mmols) in MeOH (5mL) was added K_2CO_3 (0.15 eq, 8.2 mg, 0.06 mmols) and the reaction mixture was allowed to stir overnight at room temperature [19, 22]. It was then concentrated by rotary evaporation and the residue was diluted with ethyl acetate. The mixture was then extracted with 5% HCl solution, dried with Na_2SO_4 and concentrated to yield a colorless oil 18 (85.5 mg, 67%). ^{1}H NMR (300Mz, $CDCl_3$) δ 2.21 (m, 2H), 2.04 (m, 2H), 1.79 (s, 1H), 1.48 (m, 2H), 1.37 (m, 2H), 1.20-1.12 (m, 14H).

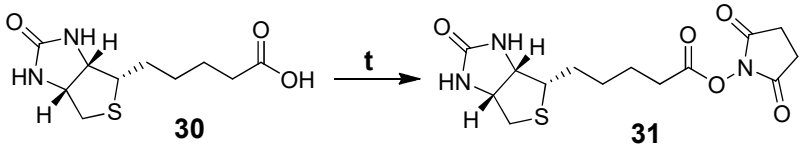
Synthesis of the tetrazine chemical reporters Tz-B and Tz-F



19 **20** (14)

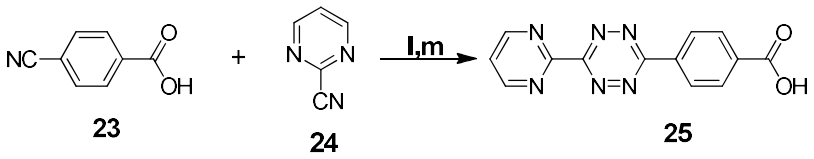
A solution of diamine **19** (5 eq, 32 mL, 0.48 mols) was prepared in 1,4 dioxane (100 mL) in a large round bottomed flask. A solution of di-tertbutyl dicarbonate (1 eq,

12.8 g, 0.06 mols) was also prepared in 1,4 dioxane (100 mL) and added to a pressure equalized graduated addition funnel. The di-tertbutyl dicarbonate solution was added to the diamine solution dropwise at room temperature and allowed to continue overnight as the initially colorless solution slowly became cloudy. The solvent was evaporated via high vacuum and the residue was resuspended in water (250 mL) the insoluble material was removed via vacuum filtration and the aqueous solution was extracted with DCM (3 x 150 mL). This was concentrated to give **20** (8 g, 69%) as a clear oil. ^{1}H NMR (300Mz, $CDCl_3$) δ 4.93 (s, 1H), 3.67 (m, 2H), 3.21 (m, 2H), 2.75 (m, 2H), 1.62 (m, 2H), 1.46 (s, 9H).



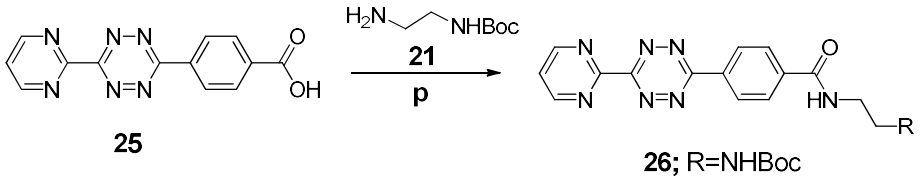
(15)

Using the method described by Kang et. al., N-Hydroxysuccinimide (1.1 eq, 0.371 g, 3.2 mmols) was added to a solution of **30** (1 eq, 1.0 g, 2.93 mmols) in DMF cooled to 0 ºC. N,N-Dicyclohexylcarbodiimide (1.3 eq, 0.786 g, 3.8 mmols) was added all at once and the reaction mixture was allowed to stir and warm to room temperature overnight [23]. The pale yellow mixture was filtered to remove any unreacted **22**, and the filtrate was concentrated to give **31** (0.873 g, 87%) as a cream-white solid. ^{1}H NMR (300Mz, DMSO) δ 6.43 (d, 2H), 4.31 (m, 1H), 4.16 (m, 1H), 2.87 (m, 2H), 2.82 (m, 4H), 2.68 (m, 2H), 1.63 (m, 3H), 1.46 (m, 3H).



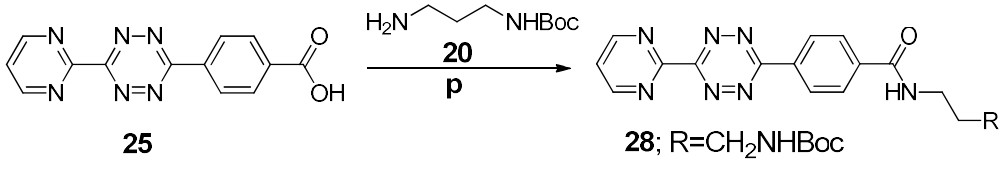
(16)

The tetrazine **(25)** was prepared as described by Chen and coworkers [16], [18], [24], [25] . To a mixture of **23** (1.5 eq, 1.96 g, 14.3 mmols) and **24** (1 eq, 0.935 g, 9.5 mmols) in dried ethanol (30 mL) was added hydrazine monohydrate (18 eq, 5.5 mL, 0.18 mols) dropwise, under inert atmosphere at room temperature. The initially grey mixture slowly changed to yellow, then burgundy as the mixture was heated to 90°C and allowed to reflux overnight. The mixture was allowed to cool and the crude product was obtained by filtering the thick mixture and washing with acetone (10 mL). The yellowish solid which resulted was placed in an erlenmeyer flask and chilled in an ice bath. Cold acetic acid (10 mL) was added to the crude product followed by a cold aqueous solution of $NaNO_2$ (2.76 g/ 10 mL H_2O). The reaction mixture changed to a bright pink-purple color and dark brown gas was observed. This mixture was allowed to stir in the ice bath for 15-20 min until no more brown gases were observed. The mixture was diluted with DCM (2 x 50 mL) and concentrated for purification. The product **25** was purified via silica gel column chromatography eluting with 95:5 hexane/ ethyl acetate to obtain the product (200 mg, 8%). ^{1}H NMR (300Mz, DMSO) δ 9.21 (d, 2H), 8.71 (d, 2H), 8.25 (d, 2H), 7.83 (m, 1H).



(17)

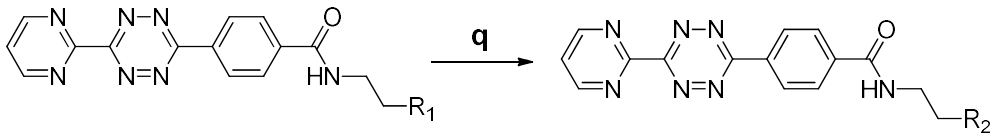
The tetrazine **(26)** was prepared as described by Chen and coworkers [16], [24]. To a solution of tetrazine **25** (1 eq, 100 mg, 0.36 mmols) and diisopropylethylamine [DIPEA] (2.5 eq, 0.233 mL, 1.34 mmols) in anhydrous DMF (10mL) was added HATU (1.5 eq, 203 mg, 0.54 mmols). This was allowed to stir at room temperature under inert atmosphere for 1 hour before addition of the protected amine 21 (1.5 eq, 86 mg, 0.54 mmols) following which the reaction mixture was allowed to stir overnight. It was then diluted with ethyl acetate (20 mL) and washed with water (3 x 20 mL), saturated sodium chloride solution (20 mL), dried over anhydrous sodium sulfate and concentrated via rotary evaporation. The crude product was purified via silica gel column chromatography eluting with 2-10% MeOH: DCM to afford the tetrazine 26, (124 mg, 82%) as a purple solid. ^{1}H NMR (300 Mz, DMSO) δ 9.24 (m, 2H), 8.71 (m, 2H), 8.16 (m, 2H), 7.86 (m, 1H), 6.85 (m, 1H), 3.03 (m, 2H), 1.69 (m, 2H), 1.39 (s, 9H).



(18)

Tetrazine **28** was prepared in a similar fashion as tetrazine **25** (1 eq, 128 mg, 0.45 mmols), DIPEA (2.5 eq, 0.199 mL, 1.14 mmols), HATU (1.5 eq, 261 mg, 0.69 mmols)

and protected amine **20** (1.5 eq, 119 mg, 0.69 mmols). After silica gel column chromatography the tetrazine **28** (119 mg, 80%) was obtained also as a purple solid. ^{1}H NMR (300Mz, $CDCl_3$) δ 9.17 (d, 2H), 8.86 (d, 2H), 8.19 (d, 2H), 7.66 (d, 1H), 4.92 (s, 1H), 3.62 (m, 2H), 3.30 (m, 2H), 1.81 (m, 2H), 1.49 (s, 9H).

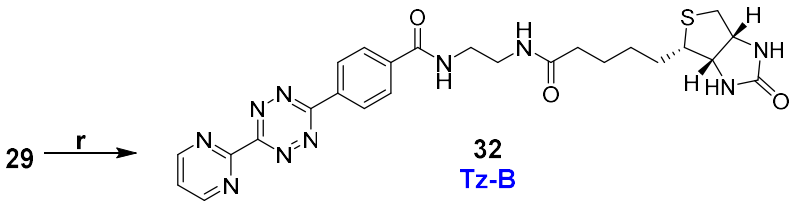


26; R_1=NHBoc
28; R_1=CH_2NHBoc

27; R_2=NH_2.HCl
29; R_2=CH_2NH_2.HCl

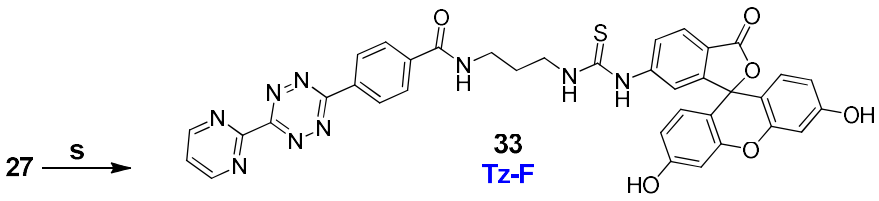
(19)

Boc-deprotection of tetrazine **26** and **28** was achieved using the methodology described by Lang et. al.[26]. The tetrazine **26** (1 eq, 124 mg, 0.29 mmols) or **28** (1 eq, 40 mg, 0.09 mmols) was dissolved in dry MeOH (~6ml or ~2 mL respectively) and chilled to 0°C. To this was added 4M HCl in dioxane (4 eq, 0.29 mL, 1.17 mmols) or (30 eq, 0.69 mL, 2.75 mmols) slowly. The reaction mixture was allowed to warm to room temperature stirring for ~1 hr following this the solvent was removed via rotary evaporation. The residue collected was suspended in diethyl ether and decanted to collect a purple solid. This was repeated and the deprotected tetrazine was collected via filtration as a purple salts, **27** (55 mg, 52%) and **29** (29 mg, 96%).



(20)

The tetrazine **29** (1 eq, 50 mg, 0.14 mmols) was dissolved in a minimum of dry DMF and DIPEA (7.6 eq, 138 µL, 1.07 mmols). The reaction mixture was stirred for 5 min before the biotin N-hydroxysuccinimide ester was added (1.3 eq, 61.8 mg, 0.18 mmols) was added. The reaction was allowed to warm to room temperature stirring overnight to yield an orange solution. The reaction mixture was concentrated and purified via silica gel column chromatography eluting with a mixture of DCM- MeOH (100:0 to 80:20) to obtain the purified tetrazine-biotin conjugate. **32** (17 mg, 33%) ^{1}H NMR (300Mz, DMSO) δ 9.17 (m, 2H), 8.79 (d, 2H), 8.12 (d, 2H), 7.79 (m, 2H), 7.19 (m, 1H), 6.85 (m, 2H), 6.55 (m, 4H), 3.80 (m, 2H), 3.50 (m, 4H), 2.10 (m, 2H), 1.38 (m, 2H).



(21)

The pink solution of tetrazine **27** (1 eq, 29 mg, 0.09 mmols) dissolved in a 2:1 mixture of dry MeOH and THF (3 mL) changed to purple upon addition of TEA (10 eq,

66 μL, 0.91 mmols). The reaction mixture was cooled to 0 ºC before fluorescein 5-isothiocyanate (FITC) (1.3 eq, 28 mg, 0.07 mmols) was added. The reaction was allowed to warm to room temperature stirring overnight to yield an orange solution. The reaction mixture was concentrated and purified via silica gel column chromatography eluting with a mixture of DCM: Acetone: MeOH (80:10:10 to 0:90:10) to obtain the purified tetrazine-fluorescein conjugate, **33** (17 mg, 33%). ^{1}H NMR (300Mz, MeOD) δ 9.19 (s, 2H), 8.66 (m, 2H), 8.14 (m, 2H), 7.95 (m, 1H), 6.41 (d, 2H), 4.31 (m, 1H), 3.43 (m, 1H), 2.83 (s, 1H), 2.78 (m, 1H), 2.10 (m, 2H), 1.57-1.29 (m, 4H), 1.34 (m,2H).

RESULTS

The products synthesized were characterized via proton NMR; the final spectra for compounds C12e, C14e, C16e, C14y, Tz-B and Tz-F are presented in Figures 7 -12.

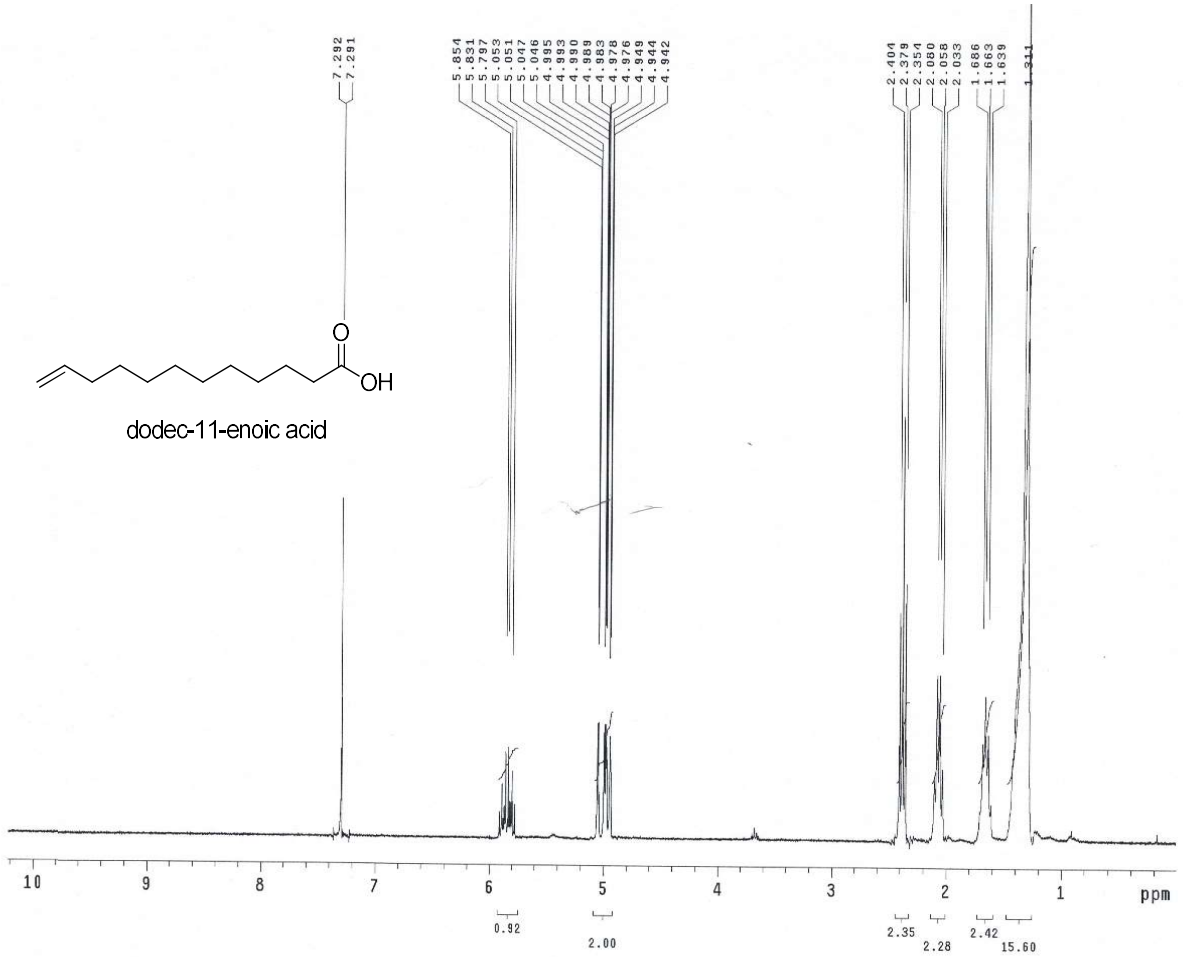


Figure 7: The proton NMR spectrum of dodec-11-enoic acid (2).

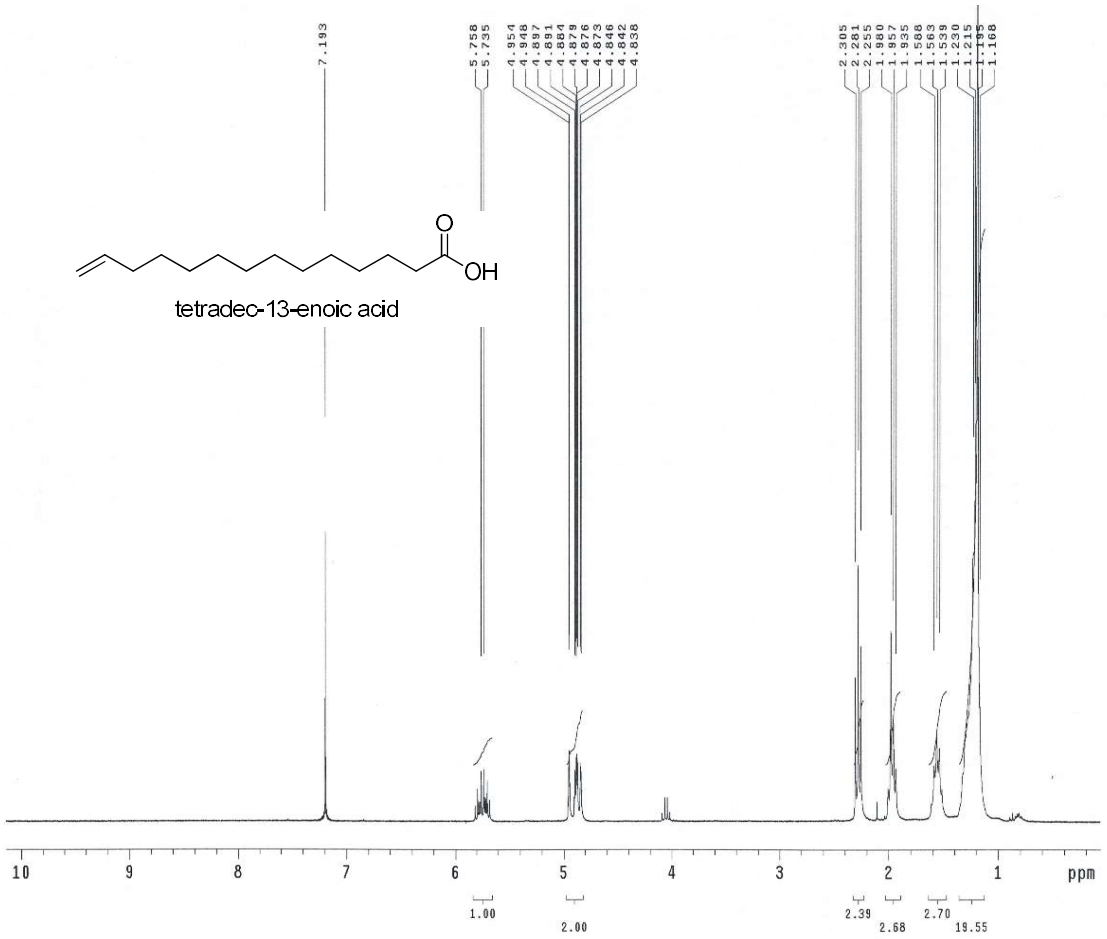


Figure 8: The proton NMR spectrum of tetradec-13-enoic acid (8).

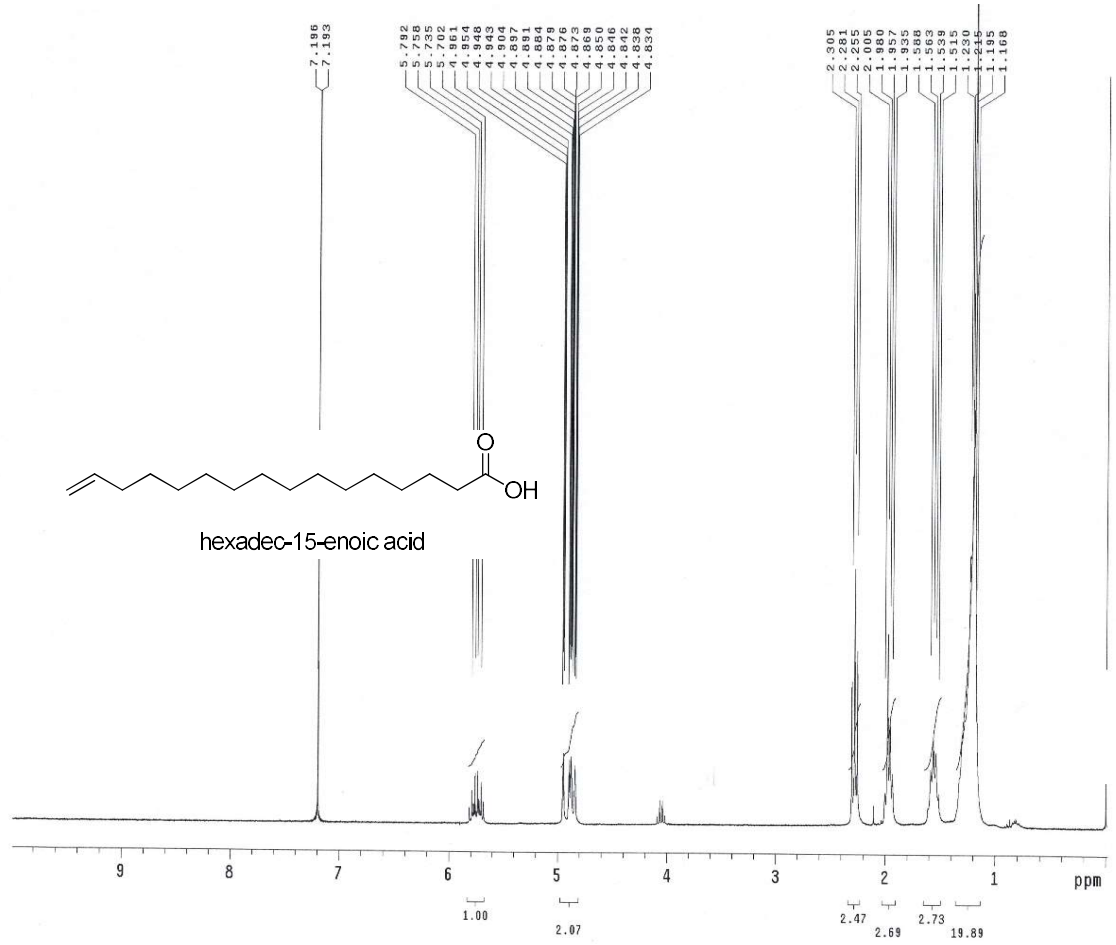


Figure 9: The proton NMR spectra of hexadec-15-enoic acid (14).

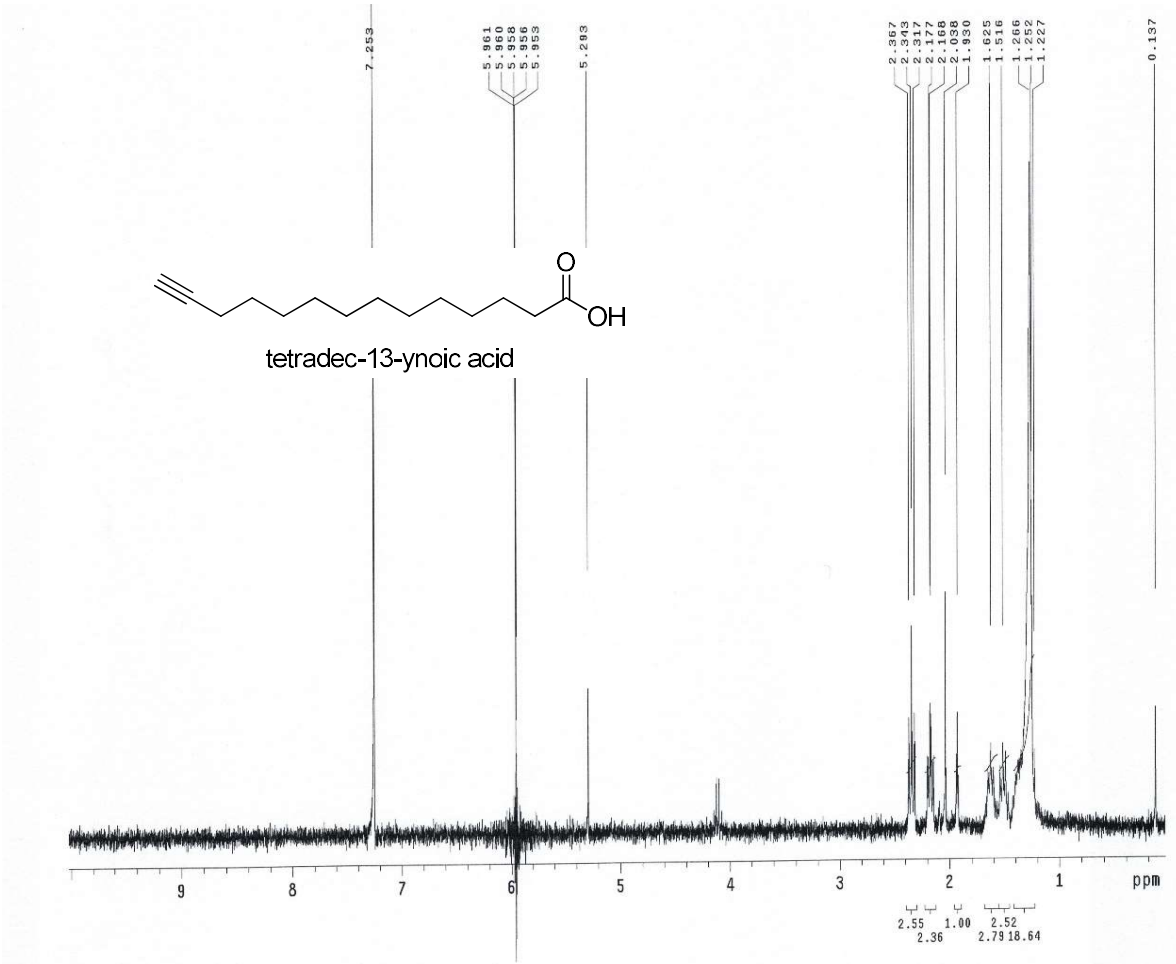


Figure 10: The proton NMR spectra of tetradec-13-ynoic acid (18).

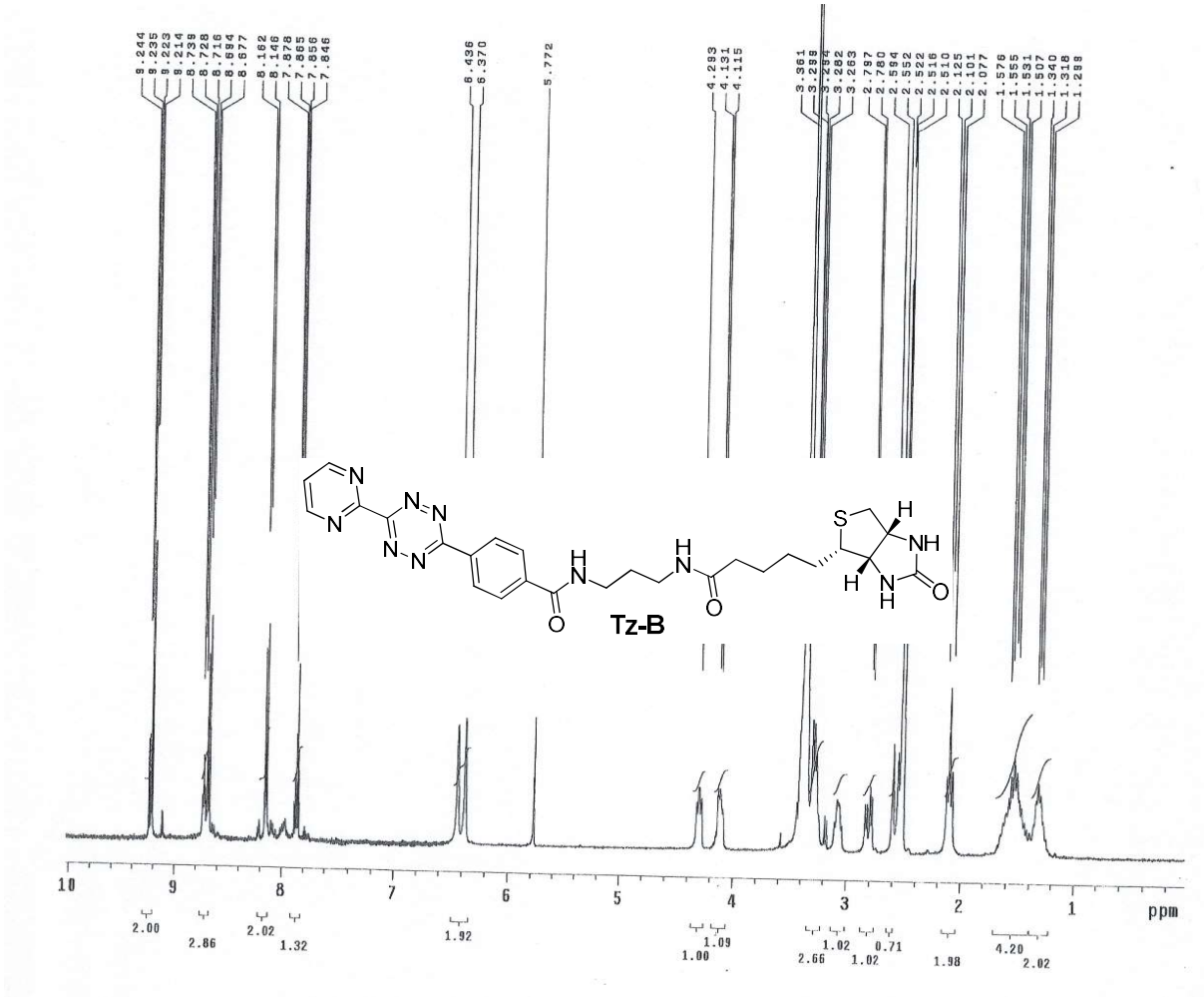


Figure 11: The proton NMR spectra of the Tetrazine- biotin conjugate probe (Tz-B) (32).

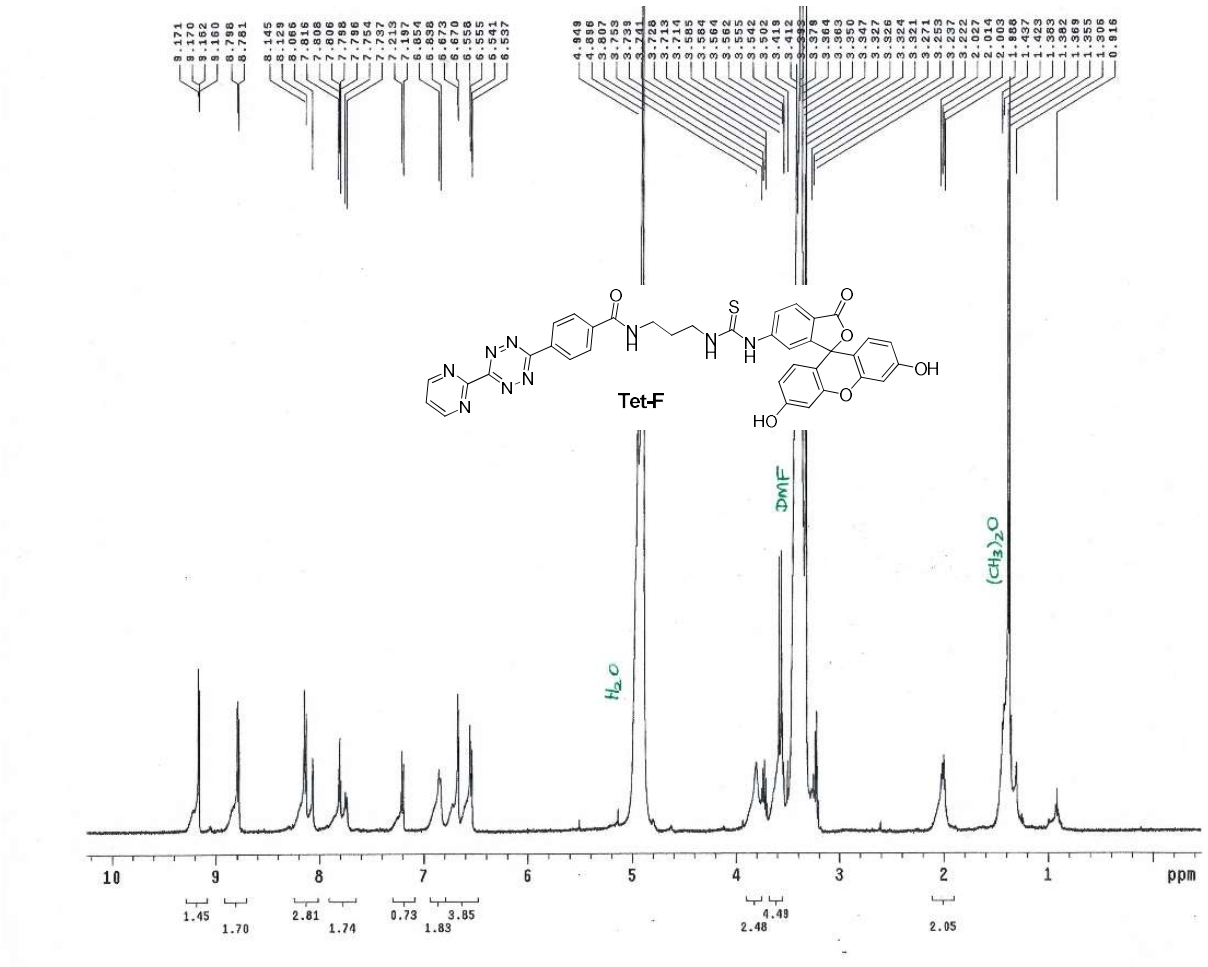


Figure 12: The proton NMR spectrum of the Tetrazine-fluorescein probe (Tz-F) (33).

THE INVERSE ELECTRON DEMAND DIELS ALDER REACTION

Bioorthogonal reactions for coupling installed chemical handles to complementary reporter molecules have become key strategies for the study of cellular modifications. However, while a variety of bioorthogonal partners have been identified only a few have found wide use in the presence of complex biological milieu. These include the Staudinger ligation and [3+2] cycloaddition between azides and alkynes[13]. In this study we demonstrate the application of the inverse electron demand Diels-Alder (iEDDA) reaction as bioorthogonal strategy for identification of fatty acylation. This reaction utilizes a terminal alkene and tetrazine as the complementary reaction partners.

REACTION MECHANISM

In contrast to the prototypical Diels Alder reaction where an electron rich diene reacts with an electron poor dienophile, in an iEDDA reaction an electron rich dienophile reacts with an electron poor diene [27]. The first and rate-determining step in the reaction mechanism is the [4+2] cycloaddition between the diene (**Tz)** and alkene dienophile to give a bicyclic adduct (**b**). This is rapidly converted, in a retro-Diels Alder step, with the production of dinitrogen to the corresponding 4, 5-dihydropyridazine (**c**), subsequent isomerization leads to the stable 1, 4-dihydro-isomer (**d**), while oxidation yields the pyridazine (**e**)[28].

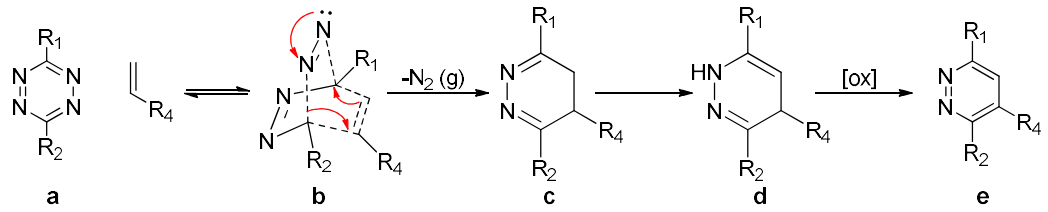


Figure 13: The iEDDA reaction between the terminal-olefin fatty acid reporter and the tetrazine capture probe. R_1 and R_2 represent the aromatic substituents of the tetrazine probe while R_4 represents the (CH_2)nCOOH of the fatty acid mimic.

REACTION KINETICS

The iEDDA reaction kinetics is governed by the separation of the HOMO $_{dienophile}$ and LUMO $_{diene}$, which allows for tuning of the reaction kinetics by varying the selection of appropriate reaction partners [28]. According to frontier orbital theory, as a general rule of thumb, electron withdrawing substituents at the 3- or 6- positions of the tetrazine lower the LUMO $_{diene}$, leading to stronger interactions with the dienophile. This promotes the crucial out of plane distortion needed to achieve the transition structure with the alkene partner and so results in an increase in the reaction rate [29].

In the case of **35a,** the absence of a strongly electron withdrawing substituent on the tetrazine or strongly electron donating substituent on the terminal alkene dienophile would suggest an unfavorable reaction. However, Lee et.al. demonstrated the favorability of this reaction in a biological system and determined the rate constant for the reaction, of **35a** and **34a**, 0.016 $M^{-1}s^{-1}$. In order to assess the suitability of **Tz-F** for biological application the reaction kinetics of this tetrazine with 6-16 carbon containing terminal alkene acids was evaluated.

The reaction kinetics was monitored under pseudo first order conditions with the dienophile (F) concentration at least 400 times greater than the diene (T) yielding a pseudo first order rate equation-

$$Rate = k' [T] \quad (22)$$

the integrated rate equation-

$$\ln[T]_t = -k't + \ln[T]_0 \quad (23)$$

With the second order rate constant given by-

$$k = k' [F] \quad (24)$$

MATERIALS AND METHODS

General information

Phosphate buffered saline dry powder (PBS), Dimethyl sulfoxide (DMSO), Hex-5-enoic acid (C6e), Oct-7-enoic acid (C8e) and Dec-9-enoic acid (C10e) were purchased from Sigma Aldrich. Dodec-11-enoic acid (C12e), Tetradec-13-enoic acid (C14e), Hexadec-15-enoic acid (C16e) were synthesized in lab. Stock solutions (1 M, 0.5 M, and 0.1 M) of the fatty alkenyl acids (C6e-C16e) were prepared in DMSO. PBS was prepared in 1 liter of micro purified water to yield a 0.1 M phosphate buffer solution (0.138 M NaCl, 0.0027 M KCl), pH 7.4 at room temperature. Stock solutions (20 mM, 0.5 mM, 0.2 mM) of the tetrazine probe Tz-F were prepared in DMSO.

Experimental procedure

Kinetic analysis of the iEDDA reaction between **Tz-F** and the terminal alkenyl fatty acids was performed using a Biotek Synergy H1 hybrid multimode micro-plate reader (Biotek, VT, USA). The terminal alkenyl fatty acids (**C6e-C16e**) were prepared in a 50% phosphate buffered saline (PBS): 50% DMSO solution at the desired concentration (2 mM – 20 mM) in individual eppendorf tubes. The **Tz-F** (5 μM) was added via micropipette and the tubes were vortexed for 30 s before being transferred to a 96 well plate. The reaction progress was monitored by the fluorescence increase at 521 nm until saturation was reached, typically 3-4 hours at 37°C, with an excitation of 493 nm. The fluorescence data versus time was fitted to a single exponential increase equation: y = m1-m2*exp (-m3*m0); using the KaleidaGraph software program such that the apparent pseudo first order reaction rate constant (k') was given by the value of m3. The second order rate constant (k) was then calculated using; $k = k'[F]$ for each terminal alkene fatty acid-tetrazine cyclo-addition reaction (Table 1).

RESULTS

In initial kinetic experiments it was observed that as the concentration of alkenyl fatty acid increased > 5mM, the anticipated fluorescence increase was observed only briefly followed by significant fluorescence quenching. This behavior was noted to be observed more drastically with the longer **C16e** fatty acid (Figure 5).

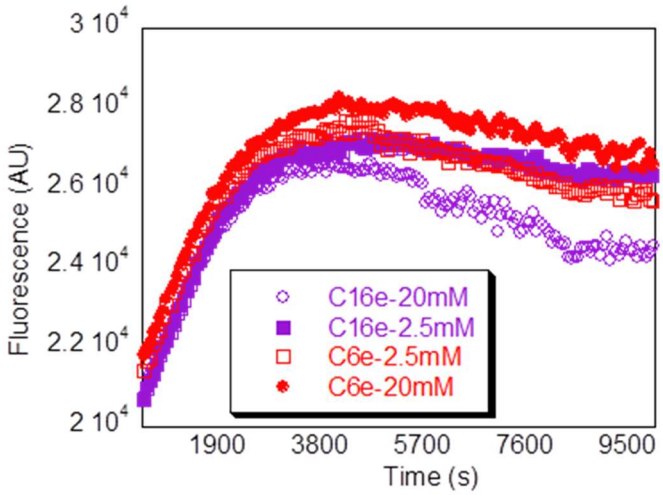


Figure 14: Shows a comparison of the fluorescence observed during the reaction of C6e and C16e (2.5 mM and 20 mM) and TzF (5 μM). Note the decrease in fluorescence observed with the reaction of Tz-F and the fatty acid C16e at a high concentration (20 mM).

This anomalous behavior is likely as a result of micelle formation during the reaction and may be compensated for by providing continuous mixing for the duration of the kinetic data acquisition. With the current experimental design using the Biotek micro-plate reader continuous mixing was not possible, as a result the observed pseudo first order rate constants were determined using concentrations between 2-3 mM alkenyl fatty acid (**C6e-C16e**) and 5μM **Tz-F** (Figure 6-8).

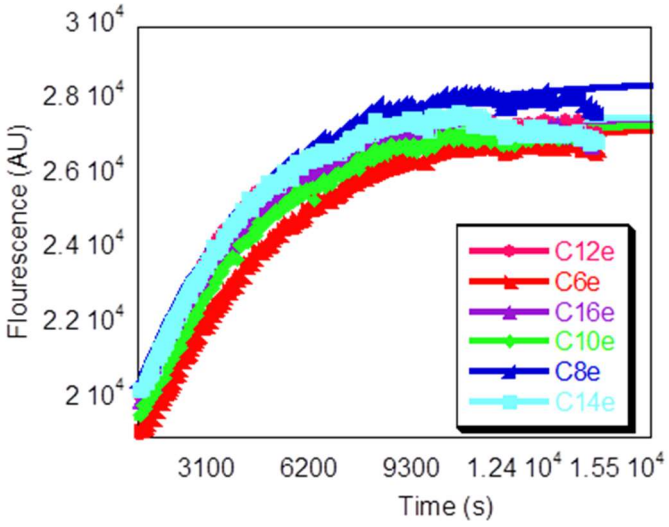


Figure 15: Shows the fluorescence increase observed for the reaction of C6e, C8e, C12e, C14e and C16e fatty acid mimics (2 mM) and Tz-F (5 μM).

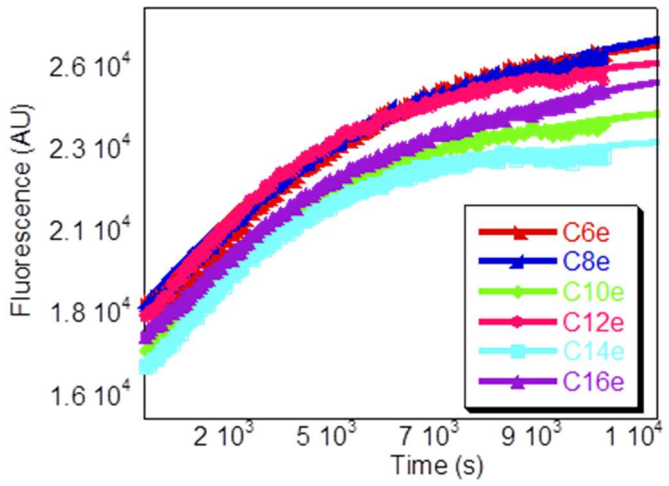


Figure 16: Shows the fluorescence increase observed for the reaction of C6e, C8e, C12e, C14e and C16e fatty acid mimics (2.5 mM) and Tz-F (5 μM).

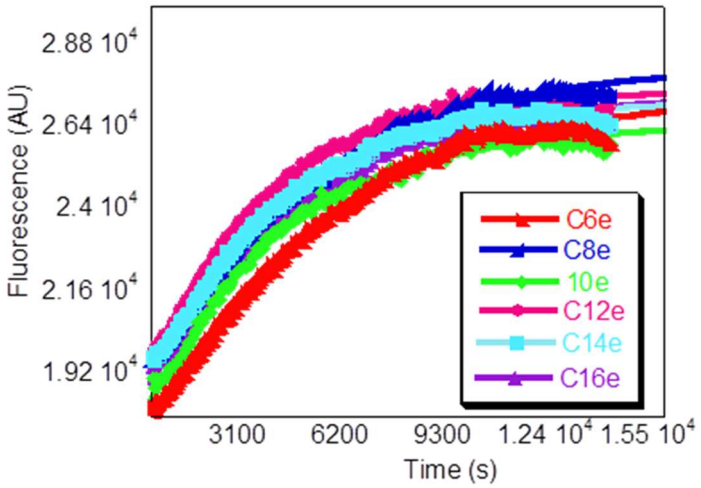


Figure 17: Shows the fluorescence increase observed for the reaction of C6e, C8e, C12e, C14e and C16e fatty acid mimics (3 mM) and Tz-F (5 μM).

Using data collected with C6e-C16e alkenyl fatty acids at 2-3μM concentrations a second order rate constant (k) was calculated (Table 1). Comparing the second order rate constant (k_{C6e}) of 0.058 $M^{-1}s^{-1}$ to that reported by Lee et. al. for hexenol of 0.036 $M^{-1}s^{-1}$ suggests that the methodology used for determination of the reaction rate constant needs to be further improved. One strategy which could be used mimics that used previously by Lee et. al. using continuous mixing during the course of data acquisition and so may eliminate complications of micelle formation with long chain fatty acid substrates during the reaction.

Table 1: Summarizes the observed reaction rate constant (k') as well as the second order rate constants (k) for the reaction of C6e-C16e with Tz-F.

Fatty Acid	k' (x 10^{-4}) s^{-1}	k $M^{-1}s^{-1}$
C6e	2.79 ± 0.059	0.48 ± 0.3
C8e	2.79 ± 0.061	0.49 ± 0.2
C10e	2.75 ± 0.064	0.56 ± 0.3
C12e	3.35 ± 0.139	0.62 ± 1.1
C14e	3.31 ± 0.152	0.67 ± 0.6
C16e	2.99 ± 0.105	0.51 ± 0.6

PRE-TARGETING AND LABELING EUKARYOTIC FATTY ACYLATION

METABOLIC LABELING OF THE EUKARYOTIC PROTEOME

Small molecules have shown to have notably better access to intracellular compartments, due to their small size [9],[30],[16]. This facilitates the installation of specific chemical epitopes onto cellular organelles by exploiting the endogenous synthesis machinery of the cell. In this metabolic labeling strategy, the terminal alkene fatty acids are introduced into the cell culture medium and taken into the cells via metabolism. The terminal alkene fatty acids may be used as a source of energy for the cells or become incorporated into cellular organelles. This provides a simplified strategy for identification of the fatty acyl modified proteins.

BIOORTHOGONAL LABELING VIA iEDDA

The installed un-natural functional group, the terminal alkene, can be targeted via ligation of a suitable complementary probe. In this instance, a biaryl tetrazine served as the complementary probe and ligation was achieved via the inverse electron demand Diels alder reaction with the terminal alkene functional group. The terminal alkene fatty acids which became incorporated during cell culture then serves as a target for the ligation of the complementary tetrazine reporter molecules [31], [32]. In this study the reactive biaryl tetrazine is conjugated to a fluorescent moiety (**Tz-F**), which facilitates the rapid visualization of the installed fatty acylated proteins by in gel fluorescence

imaging. In addition the biaryl tetrazine conjugated to biotin (**Tz-B**) was used to visualize as well as facilitate the enrichment of fatty acylated proteins for proteomic identification.

MATERIALS AND METHODS

General information

Fetal Bovine Serum (FBS), Dulbecco Modified Eagle Medium (DMEM), Phosphate Buffered Saline dry powder (PBS), Protease inhibitor cocktail, Dimethyl sulfoxide (DMSO), Trypsin, Hex-5-enoic acid (C6e), Oct-7-enoic acid (C8e) and Dec-9-enoic acid (C10e) were purchased from Sigma Aldrich. Dodec-11-enoic acid (C12e), Tetradec-13-enoic acid (C14e), Hexadec-15-enoic acid (C16e) were synthesized in lab. Compound KC6e was provided by a lab colleague, Yan Jiun Lee PhD, while compound AS1-126-1 was provided by Dr. Frederic Taran from the French Alternative Energies and Atomic Energy Commission (CEA), France.

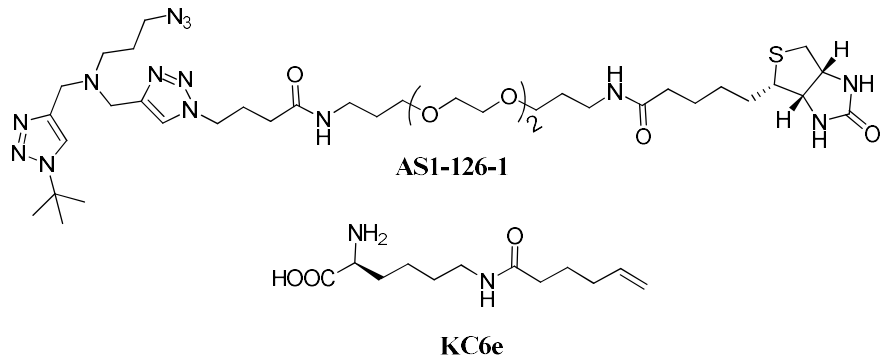


Figure 18: The structure of compound KC6e, used for the expression of sfGFP containing 6-carbon terminal alkene modified lysine residues and the chelating azide, AS1-126-1, used in labeling terminal alkyne containing proteins.

Streptavidin – HRP antibody and Pierce ECL Western Blotting Substrate were purchased through Biorad. Stock solutions (20 mM) of the fatty alkenyl acids were prepared in DMSO and filtered sequentially through 0.45uM and 0.2uM sterile filters. This stock was used to prepare a fatty acid enriched culture media by adding the necessary volume to achieve the desired concentration (typically 100 μM) to a 10:90 FBS: DMEM solution, vortexing 3-5 min and allowing the solution to rest for 8-10 mins before use.

Proteins labeled with **Tz-F** were imaged using a BioRad Molecular Imager ChemiDoc^{TM} x RS+ following SDS-PAGE. Proteins labeled with **Tz-B** were transferred to PVDF membranes using the BioRAD Trans-Blot Turbo Transfer System. After blocking with 5% BSA in PBST overnight at 4^{o}C, rinsing with PBST (10 mins) and incubation in the streptavidin-HRP antibody (1:5000) for 4 hours at room temperature (or overnight at 4^{o}C). The membrane was rinsed in PBST (10 mins) and labelled proteins visualized using Pierce^{TM} ECL western blotting substrate.

Expression of sfGFP-KC6e

A previously reported protocol was followed to genetically incorporate KC6e into sfGFP. Transformed E. coli BL21 (DE3) cells with pET-PylTsfGFPS2TAG and pEVOL-PylT-BuKRS was used to inoculate LB media and grown at 37 °C with agitation until the OD600 reached 0.6-0.8. The medium was then supplemented with 5 mM KC6e, 1 mM IPTG, and 0.2% arabinose, followed by incubation at 37 ºC with agitation for 8 h for protein expression.

After centrifugation, the cell pellet was resuspended in ice cold lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) and lysed by sonication. The lysate was centrifuged (10000 g, 4ºC, 1h) and the supernatant was collected and subjected to Ni-NTA (Qiagen) affinity chromatography. The supernatant was incubated with Ni-NTA resin (1h, 4ºC). The slurry was then loaded to a column and the protein-bound resin was washed with washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 45 mM imidazole, pH 8.0) followed by eluting the bound protein with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). The eluted protein was dialyzed (10mM ABC) and lyophilized. The lyophilized sfGFP-KC6e was dissolved in PBS and concentration determined via UV absorbance at 485 nm. [Extinction coefficient 8.33x104 M^{-1} cm^{-1}]

Incorporation of C6e-C16e fatty acids into HEK-293T cells

Passaged HEK-293T cells were seeded on 6 well plates (2-3 x 10^5 cells/mL) and grown in 10:90 FBS: DMEM (2mL) until 70-80% confluence was reached (typically 24-30 hours). The media was removed via vacuum aspiration and the cells were gently washed with PBS (2 mL) before the fatty acid enriched media was replaced and the cells grown at 37 ^oC with 5% CO_2 for 24 hours (or duration required by experiment). The enriched media was then removed by vacuum aspiration, the cells washed with PBS (2 x 1mL) and the media replaced with DMEM (2mL). The cells were returned to the incubator for an additional 2 hours. Once removed from the incubator the media was removed and the cells washed, RIPA lysis buffer (50 mM Tris-HCl at pH 7.4, 150 mM NaCl, 1 mM Na_3VO_4, 1 mM NaF, 1% Np40, protease inhibitor cocktail) was added and the cells were kept on ice for 10 mins before being scraped and collected. The cells were further lysed by 5 rounds of freeze-thaw with 30 s intervals of votexing after each thaw. The cell debris was pelleted by centrifugation 14000 rpm for 30 mins and the supernatant collected. Cell lysate protein concentration was determined using Pierce™ BCA Assay and aliquoted (100 ug protein).

Incorporation of C6y-C14y fatty acids into HEK-293T cells

Stock solutions (20 mM) of the fatty alkynl acids (C6y and C14y) were prepared in DMSO and filtered through 0.45 uM sterile filter followed by 0.2 uM sterile filter. This stock was used to prepare a fatty acid enriched culture media containing C6y (100 μM) or C14y (100 μM) in 10:90 FBS: DMEM solution. The enriched media was

vortexed 3-5 mins and allowed to rest for 8-10 mins. HEK 293T cells cultured in the presence of the alkynl fatty acids for 24 hours before being washed and lysed as described earlier.

Tetrazine labeling of sfGFP-KC6e purified proteins

The purified protein samples sfGFP-KC6e and sfGFP-wt (25 ug) were labeled with **Tz-F** (0.25 μM) or **Tz-B** (0.25 mM) in PBS buffer at pH 7.4 were incubated at 37 ºC for 4 hours. The labelled proteins were precipitated using 10% TCA, washing with acetone to remove excess of the tetrazine probe prior to electrophoresis. The precipitated proteins were re-dissolved in PBS supplemented with 2% SDS and non-reducing sample loading buffer. The sample was fully denatured by heated to 100ºC for 5 mins before being loaded to a 4-12% acrylamide gel for electrophoresis. Tz-F labeled proteins were visualized via in gel fluorescence and Tz-B labeled proteins were visualized by western blot. (Figure 10)

Tetrazine labeling of cell lysates

Cell lysates (100ug) labelled with **Tz-F** or **Tz-B** (0.25mM) in PBS buffer for 4 hours at 37ºC. The labelled proteins was precipitated using 10% TCA, washing with acetone to remove excess tetrazine probe. The precipitated proteins were re-dissolved in PBS with 2% SDS. Non-reducing sample loading buffer was added and the sample was fully denatured by heated to 100ºC for 5 mins sample before being loaded to a 4-12% acrylamide gel for electrophoresis. Proteins labeled with **Tz-F** was observed via in gel

fluorescence while those labeled with **Tz-B** was observed via western blot. (Figure 11-14)

Copper asisted azide-alkyne labelling of cell lysates

Stock solutions of the chelating azide - **AS1-126-1** (2.5 mM), $CuSO_4$ (2.5 mM) and Sodium ascorbate (125 mM) was prepared in milliQ purified water. The chelating azide (100 μM) and $CuSO_4$ (100 μM) were allowed to complex for 1 hour at room temperature. Sodium ascorbate (5 mM) was added and the solution was vortexed and allowed to stand another 10 mins before use. Cell lysates (100 ug) in PBS buffer was added to the CuAAC reagent mixture, vortexed and maintained at for 37 ºC for 4 hours for labelling. The labelled proteins was precipitated using 10% TCA, washing with acetone to remove excess tetrazine probe. The precipitated proteins were re-dissolved in PBS with 2% SDS. Non-reducing sample loading buffer was added and the sample was fully denatured by heated to 100ºC for 5 mins sample before being loaded to a 4-12% acrylamide gel for electrophoresis. Labelled proteins were observed via western blotting. (Figure 15)

HEK 293T cell fractionation

Cells cultured were cultured as described earlier in media enriched with C14e and C16e were fractionated as described by Holden et.al.[33] Enriched media was removed and the cells were washed with PBS (1 x 1 mL/well), the cells were gently dislodged using ice cold PBS (1 mL/well) and pipetted into an Eppendorf tube. The cell suspension was

centrifuged 100 x g for 10 mins at 4ºC to pellet the cells. The supernatant was removed and the cell pellet was resuspended by gentle pipetted in 400 uL Buffer 1 (150 mM NaCl, 50 mM HEPES, pH 7.4, 25 μg/mL digitonin and protease inhibitor cocktail). The cells were incubated for 10 mins end over end at 4ºC, then centrifuged at 2000 x g for 10 mins and the supernatant containing the cytosol enriched fraction was collected.

The pellet was resuspended in 400 uL Buffer 2 (150 mM NaCl, 50 mM HEPES pH 7.4 and 1% Np40) and incubated on ice for 30 mins. The cell suspension was then centrifuged 7000 x g for 10 mins. The supernatant containing the membrane bound organelles was collected. The pellet was again resuspended in 400 uL Buffer 3 (150 mM NaCl, 50 mM HEPES pH 7.4, 0.5% Sodium deoxycholate, 0.1% Sodium dodecyl sulfate and 1U/mL Benzonase). The cells were incubated overnight end over end at 4ºC followed by centrifugation 7000 x g for 10 mins. The supernatant containing the nuclear membranes and nuclear proteins were collected. To the remaining pellet was added 200 μL Buffer 4 (150 mM NaCl, 50 mM HEPES pH 7.4, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate and 100 mM dithiothretol), vortexed for 30 seconds and boiled. The sample was pelleted by centrifugation at 10000 x g for 10 mins. Protein concentration was determined for supernatant samples collected from buffers 1, 2 and 3 using the Pierce™ BCA Protein Assay Kit. Due to interference from the concentration of DTT used in buffer 4 the protein concentration was not calculated. The collected supernatant samples (100 μg protein) were labelled using Tz-F (0.25μM) for 4 hours, precipitated using 10% TCA, washed with acetone, resuspended in sample loading buffer and examined by SDS-PAGE on a 4-12% gel. (Figure 16)

RESULTS

Expressed and purified sfGFP with a C6e modification was labelled with Tz-F or Tz-B to verify the functionality of the terminal alkene-tetrazine labeling system (Figure 10), as illustrated the terminal alkene containing sfGFP-KC6e was labelled exclusively confirming the selectivity and specificity of this reaction.

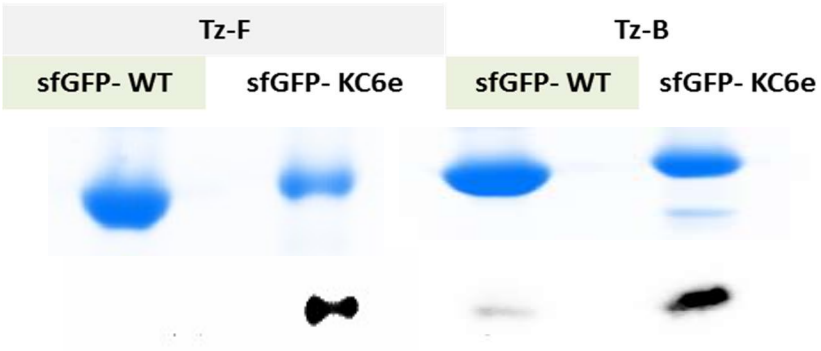


Figure 19: sfGFP-WT and sfGFP-KC6e (25 ug) were labeled with 0.25 mM Tet-F and Tet-B for 6 hrs at 37ºC and visualized via in gel fluorescence and western blot analysis respectively.

An initial test of the metabolic labeling conditions using fatty acid enriched media with C6e (100 μM) or C14e (100 μM) incubating for 24 hours followed by cell lysis and labeling of whole cell lysates with Tz-F or Tz-B indicated the presence of fatty acylated proteins [30]. (Figure 20)

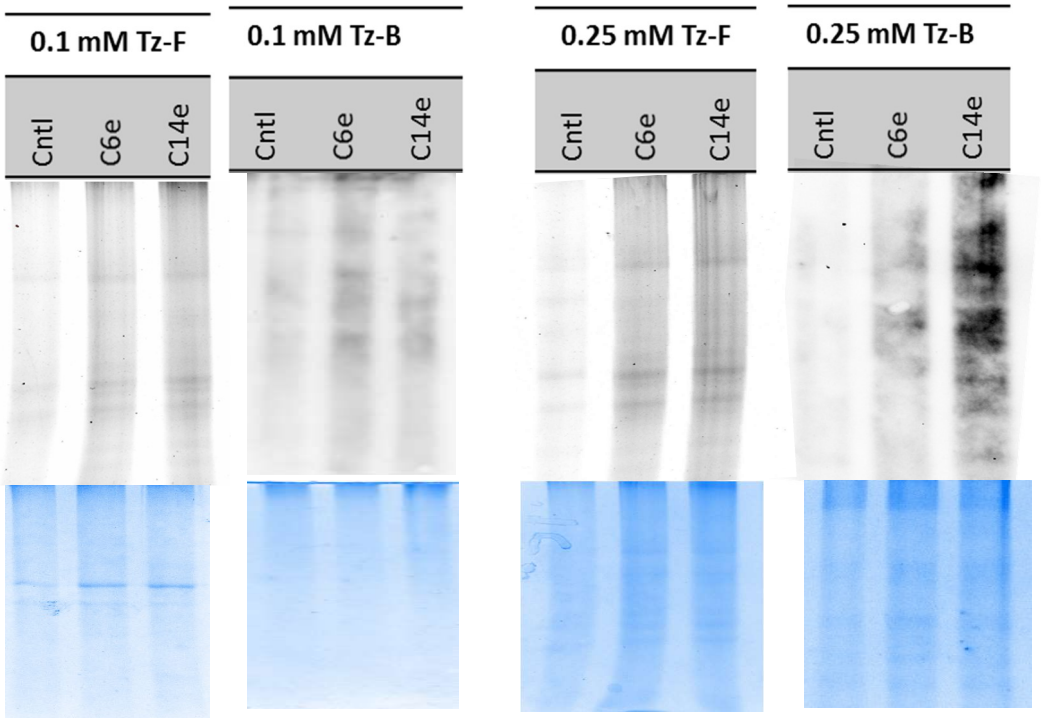


Figure 20: In gel fluorescence and chemiluminescent western blot of HEK 293T cell lysate from cells cultured in the presence of C6e or C14e (100 μM) and labelled with Tz-F (0.1 mM, 0.25 mM) or Tz-B (0.1 mM, 0.25 mM) for 4 hours. For western blots the proteins were transferred to PVDF membranes, blocked with 5% BSA, incubated in Streptavidin-HRP and visualized.

In order to identify the best conditions for achieving optimal metabolic labeling, various concentrations of the different terminal alkene fatty acids were provided in the growth medium of HEK-293T cells. Using both the level of labelled proteins observed via western blot as well as the appearance of the cultured cells it was determined that 100 uM of fatty acid was sufficient to gain the best results. While higher concentrations appeared to give increased incorporation of terminal alkene the level of cell confluence reached following 24 hours of culture suggested that higher concentrations also resulted in increased cellular stress leading to decreased growth.

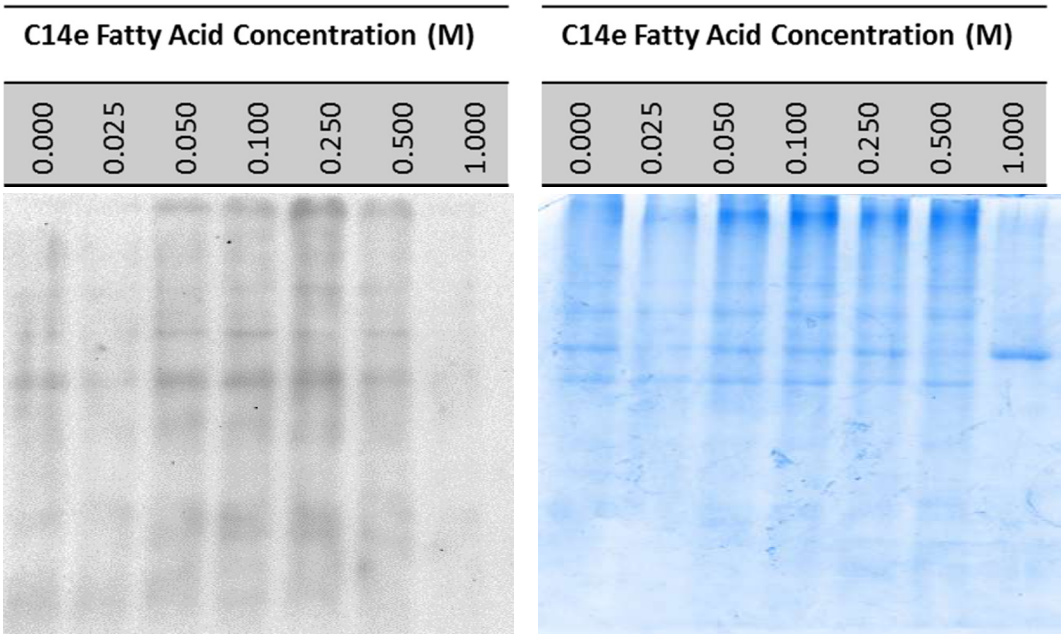


Figure 21: Chemiluminescence observed via western blotting of cell lysates cultured in varying concentrations of C14e enriched media to determine the optimum concentration for metabolic labeling. The cell lysates were labelled with Tz-B (0.25 mM) for 4 hours, transferred to PVDF membranes, blocked with 5% BSA, incubated in Streptavidin-HRP and visualized.

Metabolic labelling conditions were further optimized by determining the best duration for culturing cells in the fatty acid enriched media. An enriched media containing C14e (100 μM) was prepared and cells were grown for 6, 12, 18, 24, 36 and 48 hours. The enriched media was replaced after 24 hours to ensure that the growth media was not toxic. The control sample was cultured for 36 hours in 10:90 FBS/DMEM with DMSO. (Figure 22)

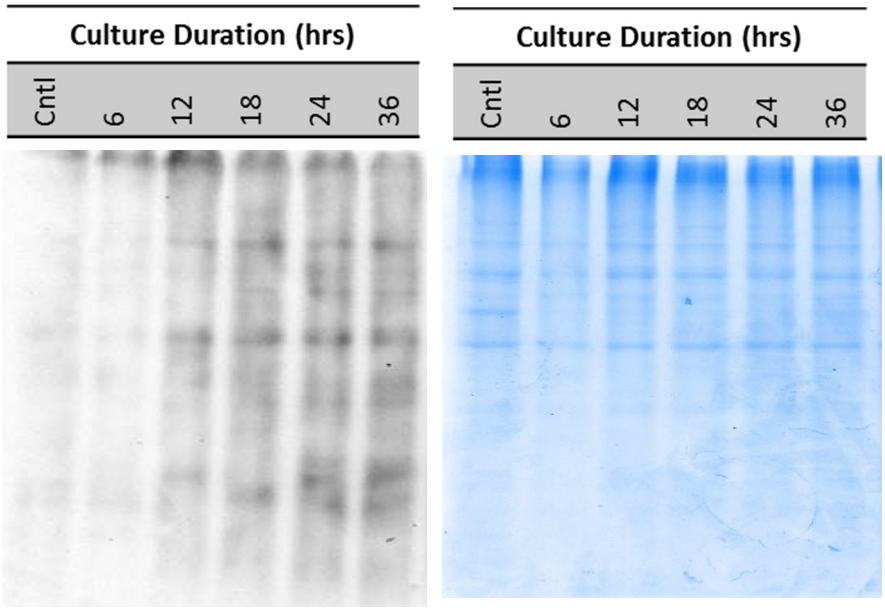


Figure 22: Chemiluminescence observed via western blotting of cell lysates; cells were cultured in C14e (100 μM) enriched media for varying durations 6, 12, 18, 24 and 36 hours. The cell lysates were labeled with Tz-B (0.25 mM) for 4 hours, transferred to PVDF membranes, blocked with 5% BSA, incubated in Streptavidin-HRP and visualized.

Cells were also cultured for 24 hours in media enriched with fatty acids of varying chain lengths C6e, C8e, C10e, C12e, C14e, and C16e. A control sample was cultured for 24 hours in 10:90 FBS/DMEM with DMSO. The results suggest that there is a preference for incorporation of longer chained fatty acid substrates by the increased labeling observed in the C14e-C16e lysates. This result was anticipated as myristoyl and palmitoyl fatty acid modifications are widely identified modifications in the eukaryotic proteome. (Figure 23)

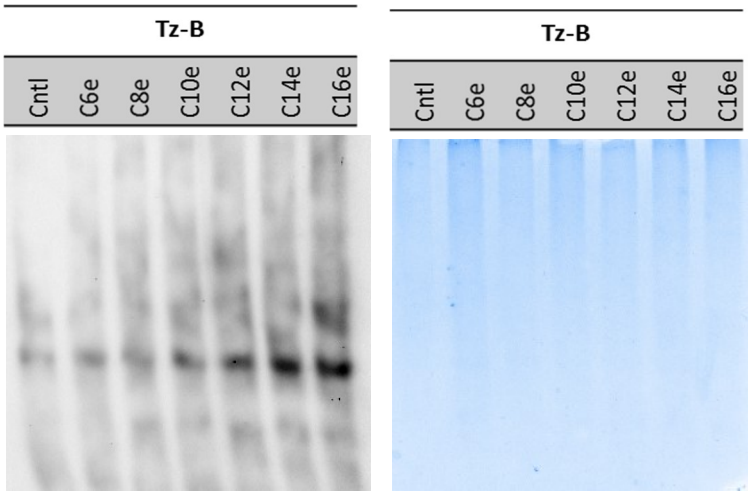


Figure 23: Chemiluminescent visualization of modified cell lysates cultured in enriched media with 100 μM C6e- C16e for 24 hrs. The collected cell lysates were labelled with Tz-B (0.25 mM) for 4 hours , transferred to PVDF membranes, blocked with 5% BSA, incubated in Streptavidin-HRP.

To investigate the difference in fatty acylation targets between a terminal alkene fatty acid (C6e, C14e) and terminal alkyne fatty acid (C6y, C14y) HEK cells were cultured in media supplemented with the respective fatty acid (100 μM) for 24 hours. The cell lysates were labelled using either Tz-B (0.1 mM) or a chelating biotin azide AS1-126-1 (0.1 mM) for 4 hours. The results though preliminary indicate that the terminal alkyne fatty acid modifies different proteins than the terminal alkene fatty acid. Also the terminal alkyne fatty acid has a broader range of targets than the terminal alkene which may account for the number of falsely identified fatty acid modification sites. (Figure 24)

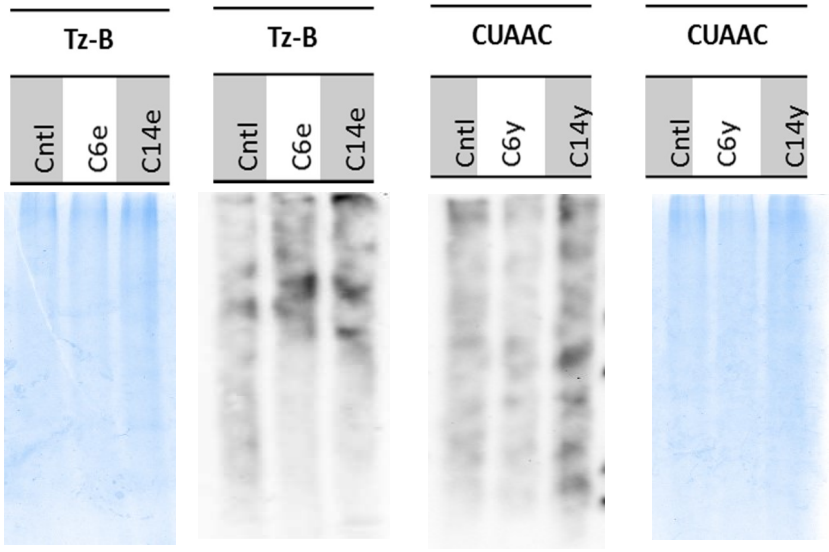


Figure 24: Chemiluminescent visualization of cell lysates obtained from cells cultured in enriched media with 100 μM C6e, C6y, C14e or C14y for 24 hrs. The collected cell lysates from C6e and C14e treatments were labelled with Tz-B (0.1 mM) for 4 hours. The cell lysates treated with C6y and C14y were labelled with AS1-126-1 (0.1 mM), $CuSO_4$ (0.1 mM) and sodium ascorbate (5 mM) for 4 hours. The modified proteins were then transferred to PVDF membranes, probed with Strep-HRP and visualized.

One strategy which was attempted to deconvolute the complex mixture of proteins obtained following cell lysis uses cellular fractionation. (Figure 25) By methodically separating the cell lysates into its cytosolic proteins (B1), proteins contained in the cells membranous organelles (B2) including the endoplasmic reticulum, golgi, mitochondria and nuclear luminal proteins as well as those in the nucleus and nuclear membrane (B3) and insoluble proteins (B4) it would be possible to evaluate the localization of the fatty acyl modifications as well as reduce the sample complexity.

This process would be very advantageous for mass spectroscopy analysis which often suffers due to the difficulties associated with deconvolution of samples prior to ionization. In our experiments, the results indicate that fatty acyl modifications are

focused mainly on membranous cellular organelles. This aligns with findings which have identified fatty acyl modifications on cell membranes and other hydrophobic proteins[34], [7].

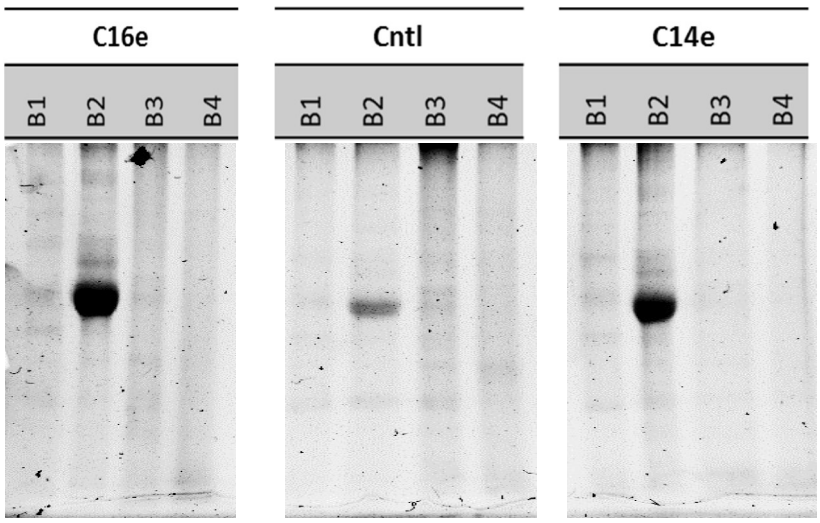


Figure 25: In gel fluorescence of cell lysates obtained from cells cultured in enriched media with 100 μM C14e or C16e for 24 hrs. The cell lysates were collected and labelled with Tz-F (0.25 mM) for 4 hours and the modified proteins visualized by in gel fluorescence.

CONCLUSIONS AND FUTURE PERSPECTIVES

In using a tetrazine- terminal olefin methodology we have demonstrated that a terminal alkene fatty acid mimic can be incorporated successfully into the eukaryotic proteome. Using the tetrazine fluorescein and biotin conjugated probes we have selectively labelled fatty acid modified cellular components. Evidence has also been obtained to show that fatty acyl modifications are mainly localized on the cells membranous organelles including the endoplasmic reticulum, golgi, mitochondria and nuclear luminal proteins.

A comparison of the labelling of similar terminal alkene and alkyne fatty acid surrogates provides insufficient results to suggest that the iEDDA strategy is more efficient or effective than CuAAC in identifying fatty acid modifications. However, with further work this alternative labeling strategy may be advantageous in validating some of the catalogued sites of fatty acyl modifications. This tetrazine-alkene system would provide another unique data set which may potentially be more representative of endogenous processes than those currently being used. Furthermore, the simplicity of this chemical strategy makes it readily adapted to fit other research objectives offering another novel research tool to the extended scientific community.

REFERENCES

[1] C. T. Walsh, S. Garneau‐Tsodikova, G. J. Gatto, *Angewandte Chemie International Edition* **2005**, *44*, 7342-7372.

[2] J. Seo, K.-J. Lee, *J. Biochem. Mol. Biol.* **2004**, *37*, 35-44.

[3] X. Gao, R. N. Hannoush, *Nat. Chem. Biol.* **2014**, *10*, 61-68.

[4] D. R. Colquhoun, A. E. Lyashkov, C. U. Mohien, V. N. Aquino, B. T. Bullock, D. R. M. Graham, R. R. Dinglasan, B. J. Agnew, *Proteomics* **2015**, *15*, 2066-2077.

[5] N. FAEligrgeman, J. Knudsen, *Biochem. J* **1997**, *323*, 1-12.

[6] C. Arrowsmith, J. E. Audia, C. Austin, J. Baell, J. Bennett, J. Blagg, C. Bountra, P. E. Brennan, P. J. Brown, M. E. Bunnage, C. Buser-Doepner, R. M. Campbell, A. J. Carter, P. Cohen, R. A. Copeland, B. Cravatt, J. L. Dahlin, D. Dhanak, A. M. Edwards, M. Frederiksen, *Nat. Chem. Biol.* **2015**, *11*, 536-541.

[7] T. Peng, E. Thinon, H. C. Hang, *Curr. Opin. Chem. Biol.* **2016**, *30*, 77-86.

[8] E. M. Storck, R. A. Serwa, E. W. Tate, *Biochem. Soc. Trans.* **2013**, *41*, 56-61.

[9] R. N. Hannoush, J. Sun, *Nat. Chem. Biol.* **2010**, *6*, 498-506.

[10] G. Charron, J. Wilson, H. C. Hang, *Curr. Opin. Chem. Biol.* **2009**, *13*, 382-391.

[11] K. N. Chuh, M. R. Pratt, *Curr. Opin. Chem. Biol.* **2015**, *24*, 27-37.

[12] M. Yu, L. P. S. de Carvalho, G. Sun, J. S. Blanchard, *J. Am. Chem. Soc.* **2006**, *128*, 15356-15357.

[13] N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297-2299.

[14] Y.-Y. Yang, J. M. Ascano, H. C. Hang, *J. Am. Chem. Soc.* **2010**, *132*, 3640-3643.

[15] M. D. Resh, *Nat. Chem. Biol.* **2006**, *2*, 584-590.

[16] Y.-J. Lee, Y. Kurra, Y. Yang, J. Torres-Kolbus, A. Deiters, W. R. Liu, *Chem. Commun. (Cambridge, U. K.)* **2014**, *50*, 13085-13088.

[17] R. Pipkorn, W. Waldeck, B. Didinger, M. Koch, G. Mueller, M. Wiessler, K. Braun, *J. Pept. Sci.* **2009**, *15*, 235-241.

[18] M. R. Karver, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2011**, *22*, 2263-2270.

[19] A. A. Wube, A. Hüfner, C. Thomaschitz, M. Blunder, M. Kollroser, R. Bauer, F. Bucar, *Bioorg. Med. Chem.* **2011**, *19*, 567-579.

[20] F. Guillod, J. Greiner, J. G. Riess, *Chem. Phys. Lipids* **1995**, *78*, 149-162.

[21] T. Řezanka, L. Siristova, K. Melzoch, K. Sigler, *Chem. Phys. Lipids* **2009**, *159*, 104-113.

[22] W. P. Heal, S. R. Wickramasinghe, R. J. Leatherbarrow, E. W. Tate, *Org. Biomol. Chem.* **2008**, *6*, 2308-2315.

[23] S. Kang, L. Mou, W. J. Brouillette, P. E. Prevelige, *Rapid communications in mass spectrometry : RCM* **2009**, *23*, 1719-1726.

[24] W. Chen, D. Wang, C. Dai, D. Hamelberg, B. Wang, *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 1736-1738.

[25] H. S. Beckmann, A. Niederwieser, M. Wiessler, V. Wittmann, *Chemistry–A European Journal* **2012**, *18*, 6548-6554.

[26] K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters, J. W. Chin, *Nature Chem.* **2012**, *4*, 298-304.

[27] K. Lang, J. W. Chin, *Chem. Rev. (Washington, DC, U. S.)* **2014**, *114*, 4764-4806.

[28] A.-C. Knall, C. Slugovc, *Chem. Soc. Rev.* **2013**, *42*, 5131-5142.

[29] F. Liu, Y. Liang, K. Houk, *J. Am. Chem. Soc.* **2014**, *136*, 11483-11493.

[30] R. N. Hannoush, N. Arenas-Ramirez, *ACS Chem. Biol.* **2009**, *4*, 581-587.

[31] G. C. Rudolf, W. Heydenreuter, S. A. Sieber, *Curr. Opin. Chem. Biol.* **2013**, *17*, 110-117.

[32] Y. Zhao, O. N. Jensen, *Proteomics* **2009**, *9*, 4632-4641.

[33] P. Holden, W. A. Horton, *BMC research notes* **2009**, *2*, 243.

[34] A. F. Roth, J. Wan, W. N. Green, J. R. Yates, N. G. Davis, *Methods* **2006**, *40*, 135-142.

APPENDIX SUPPLEMENTARY FIGURES

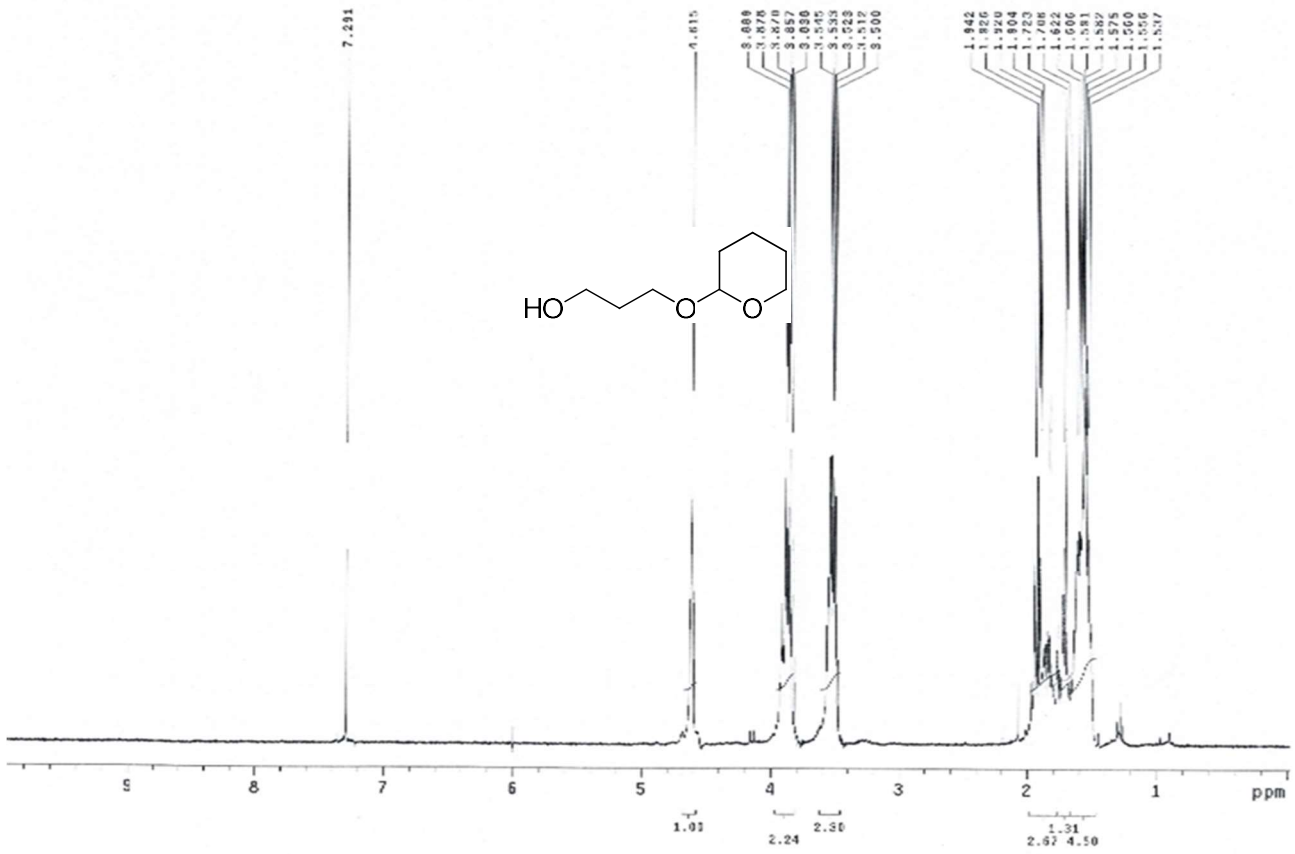


Figure S1: The proton NMR spectra of compound 4.

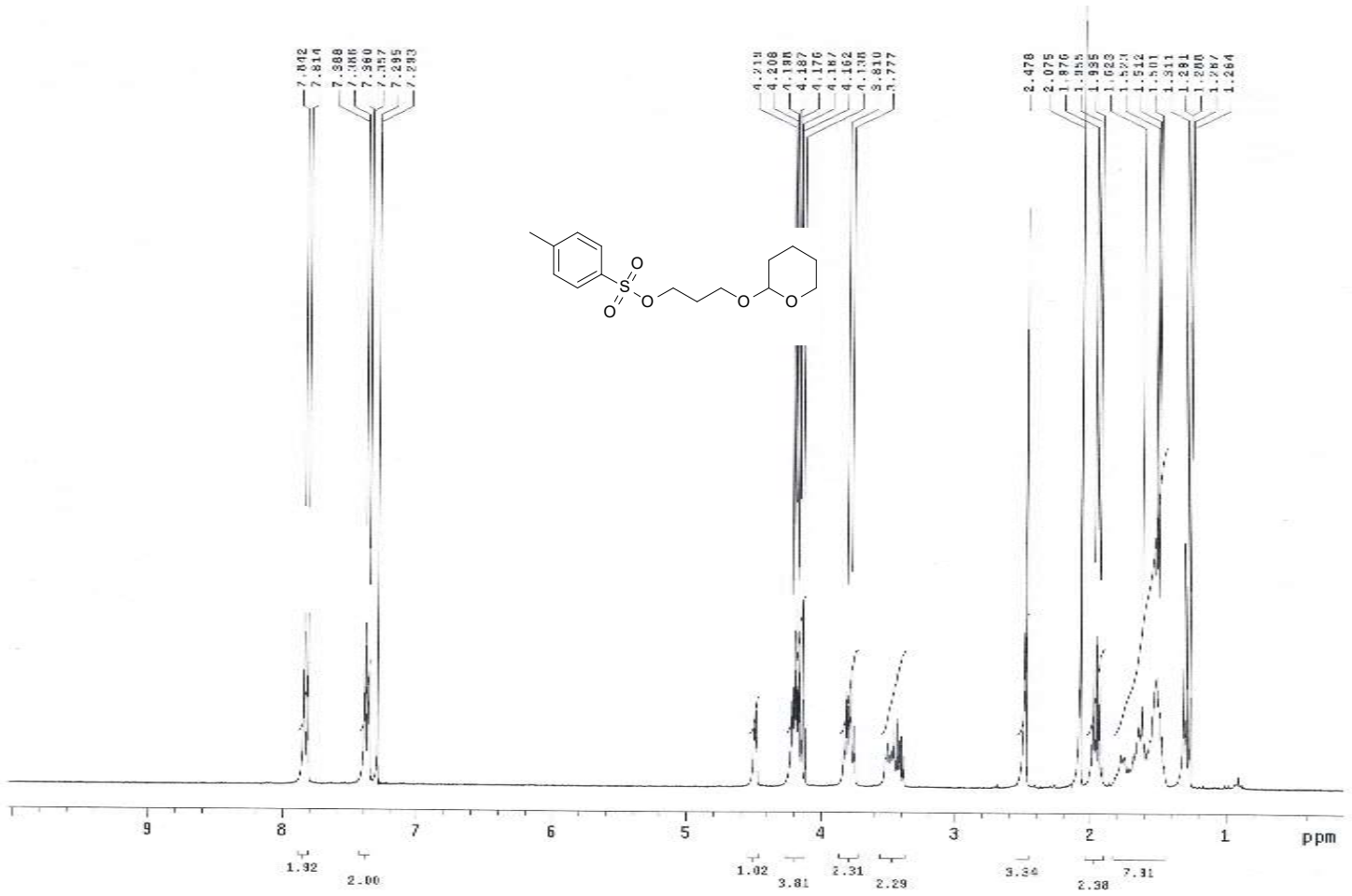


Figure S2: The proton NMR spectra of compound 5.

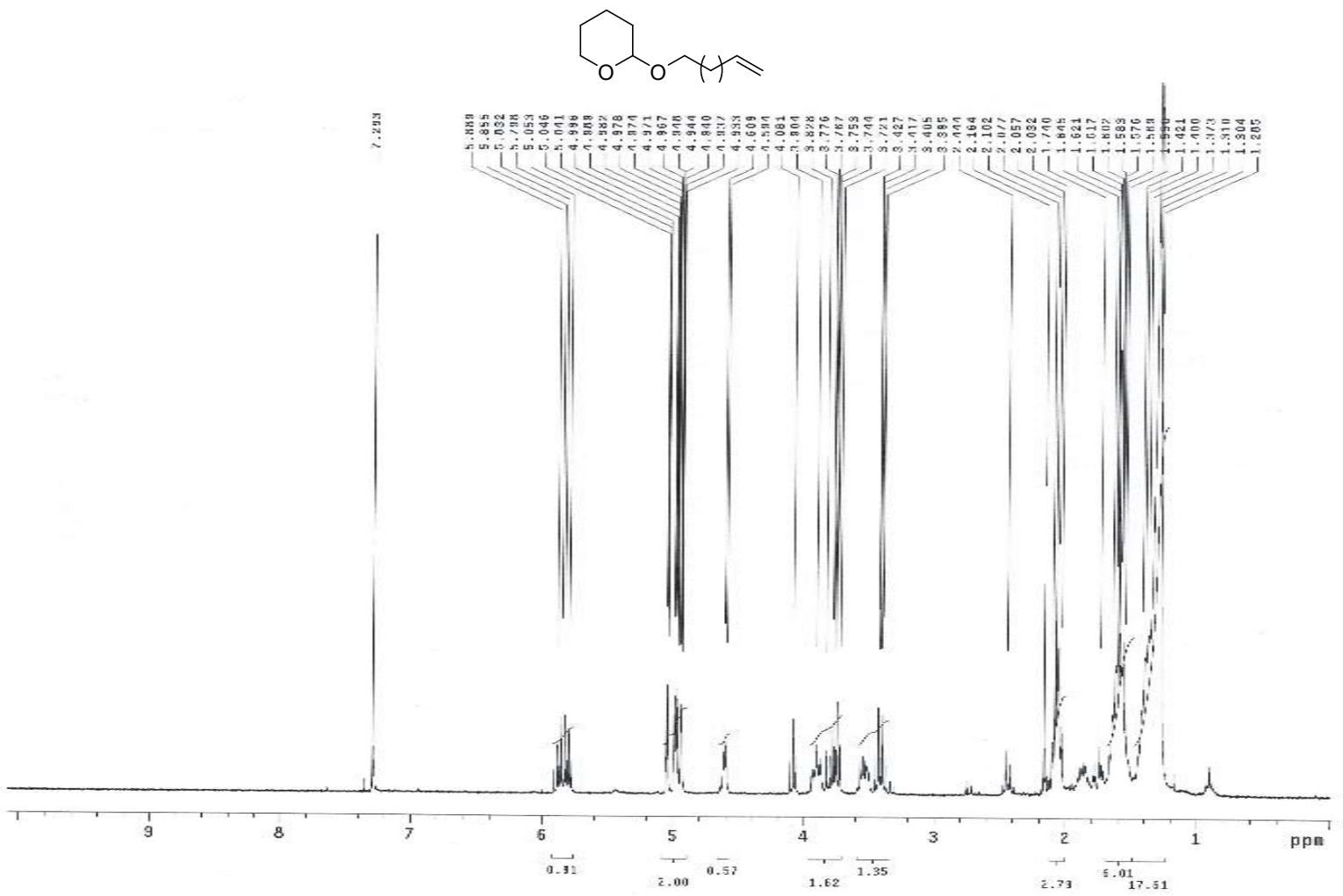


Figure S3: The proton NMR spectra of compound 6.

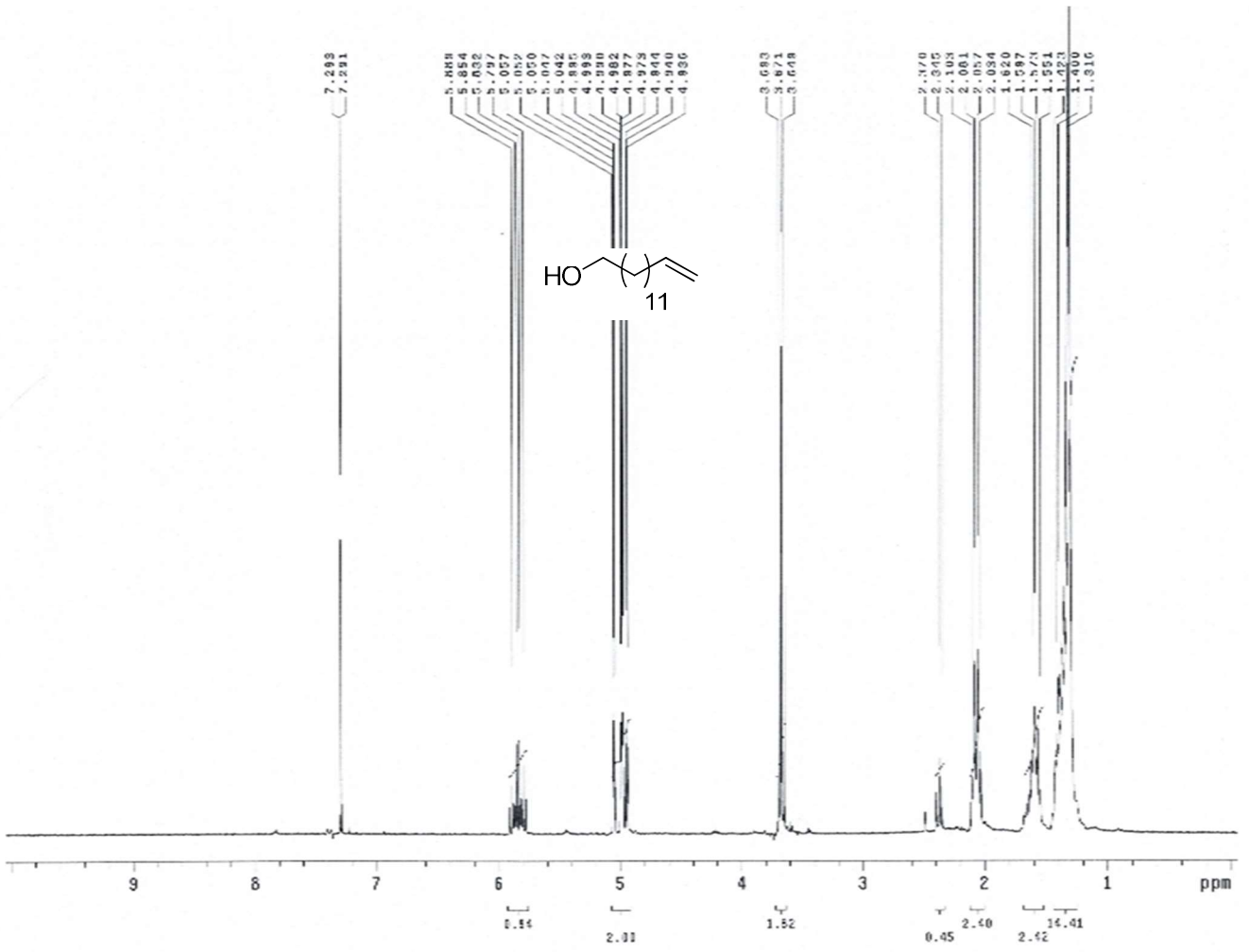


Figure S4: The proton NMR spectra of compound 7.

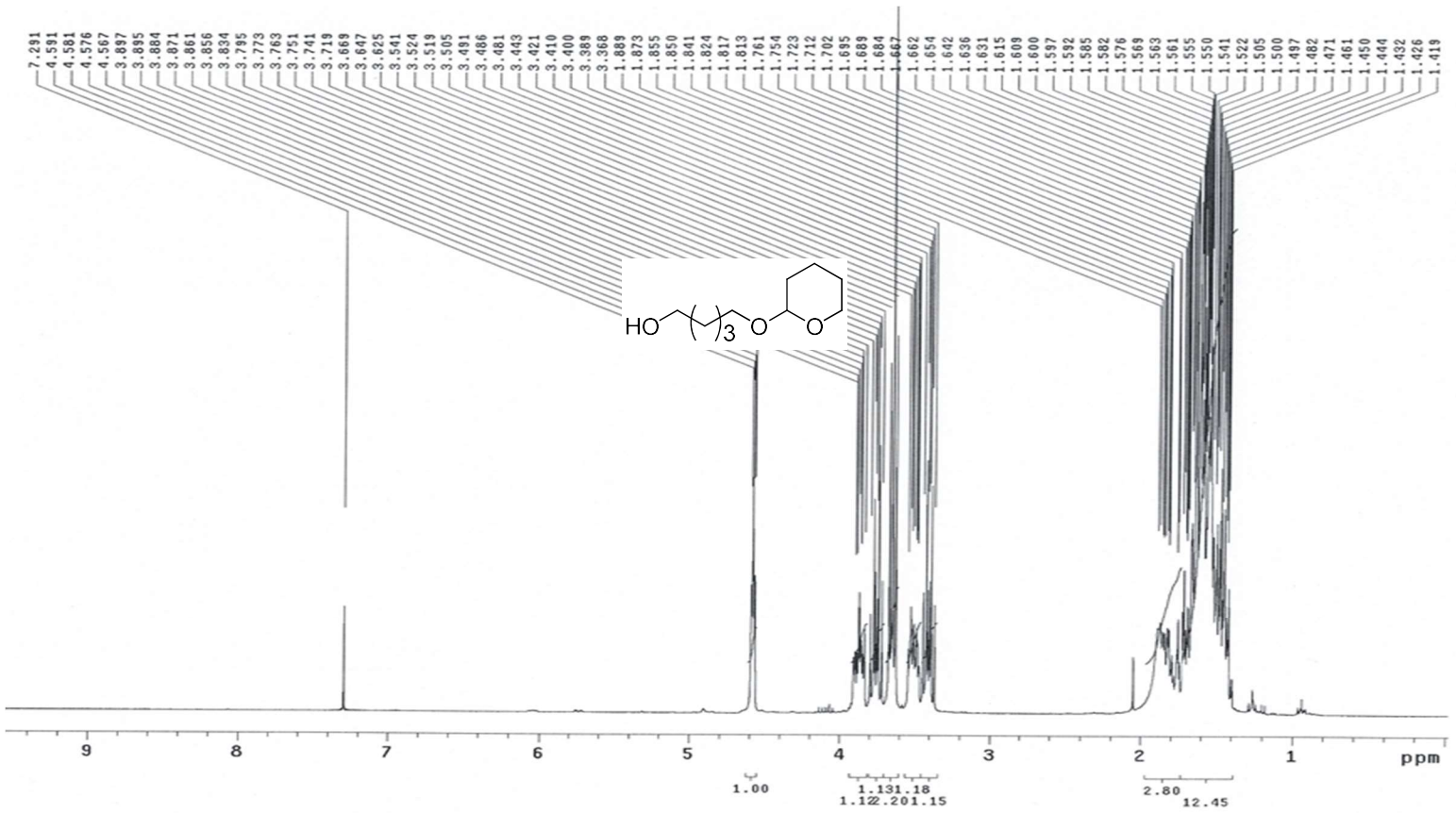


Figure S5: The proton NMR spectra of compound 10.

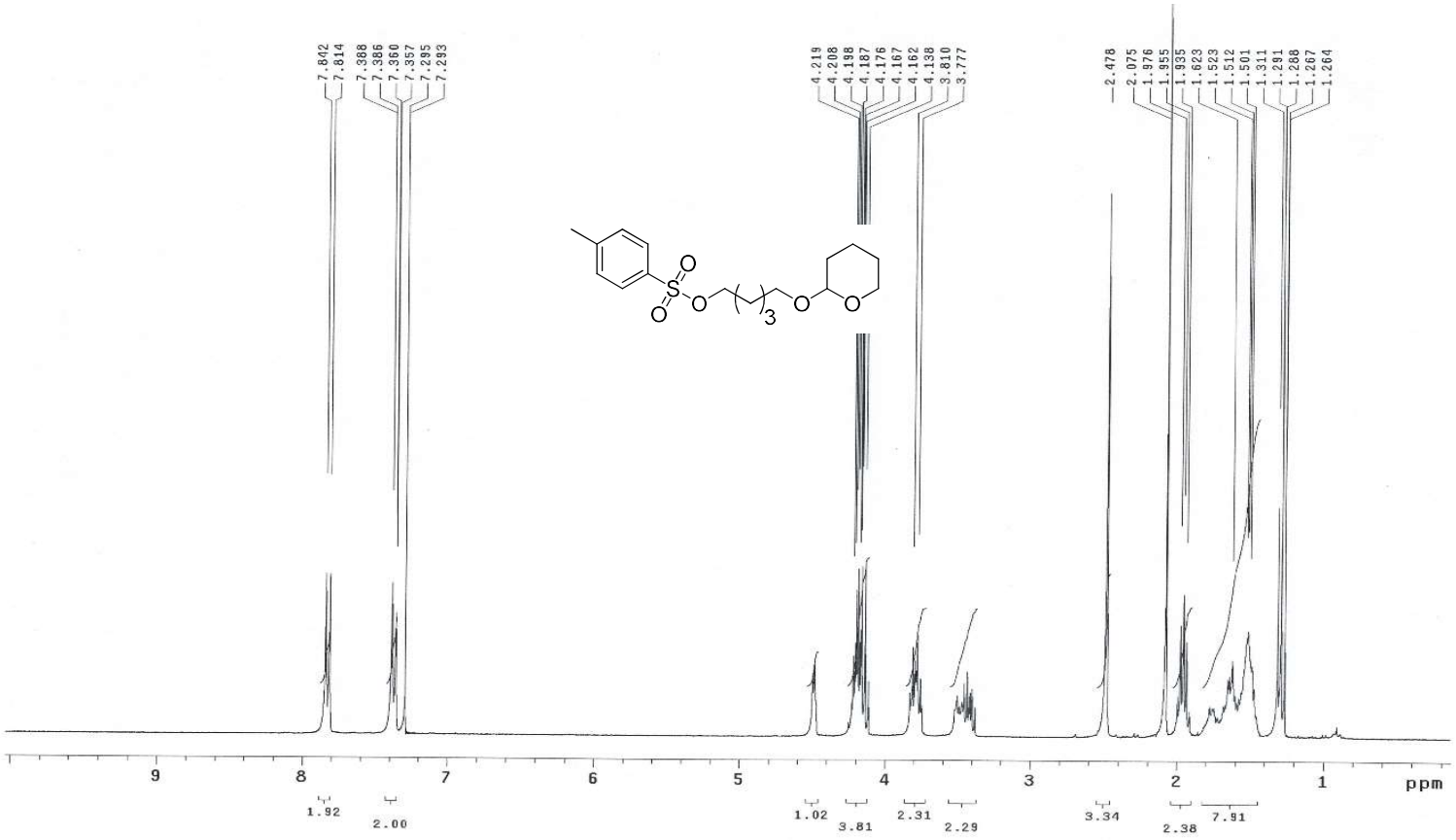


Figure S6: The proton NMR spectra of compound 11.

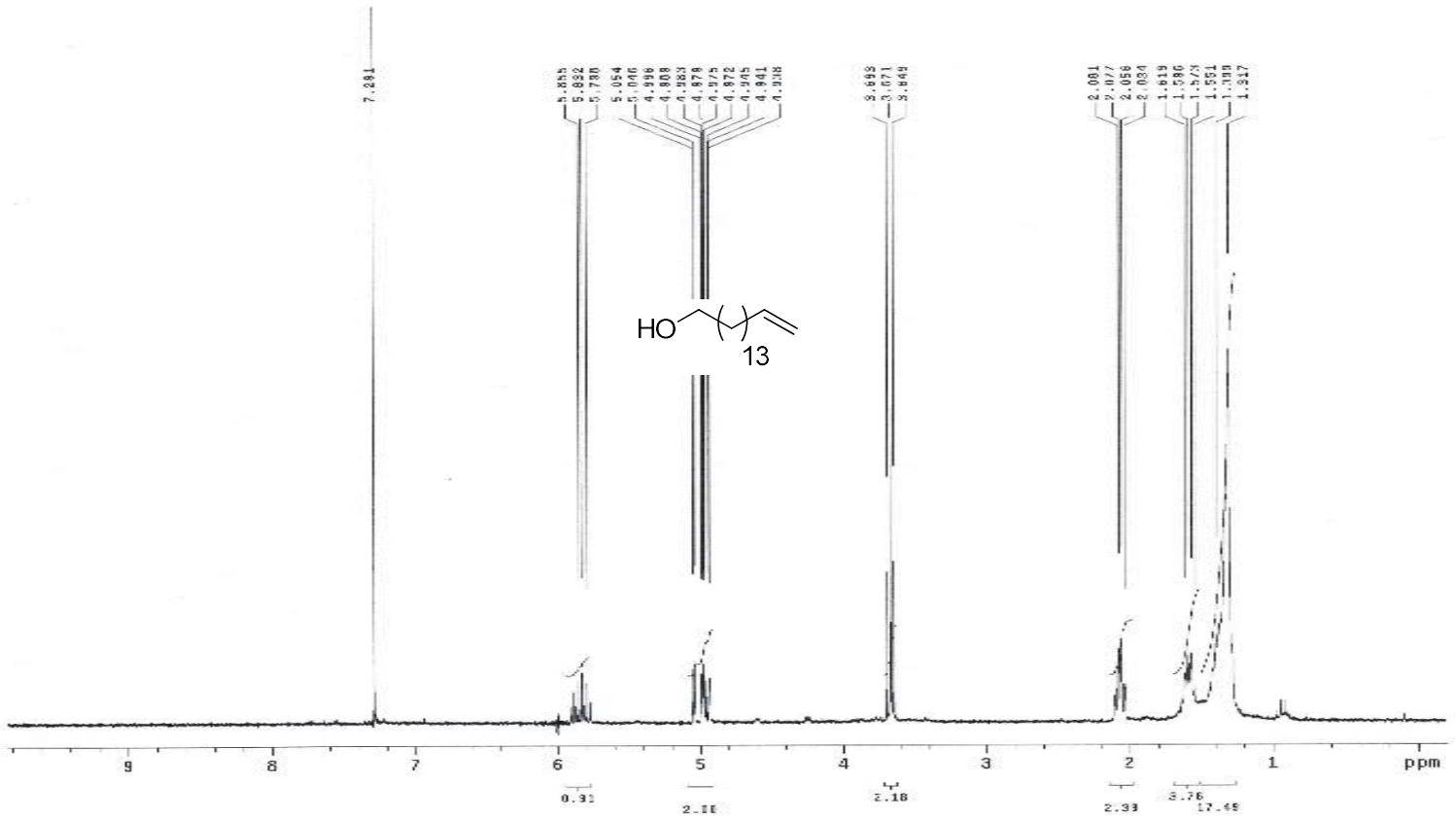


Figure S7: The proton NMR spectra of compound 13.

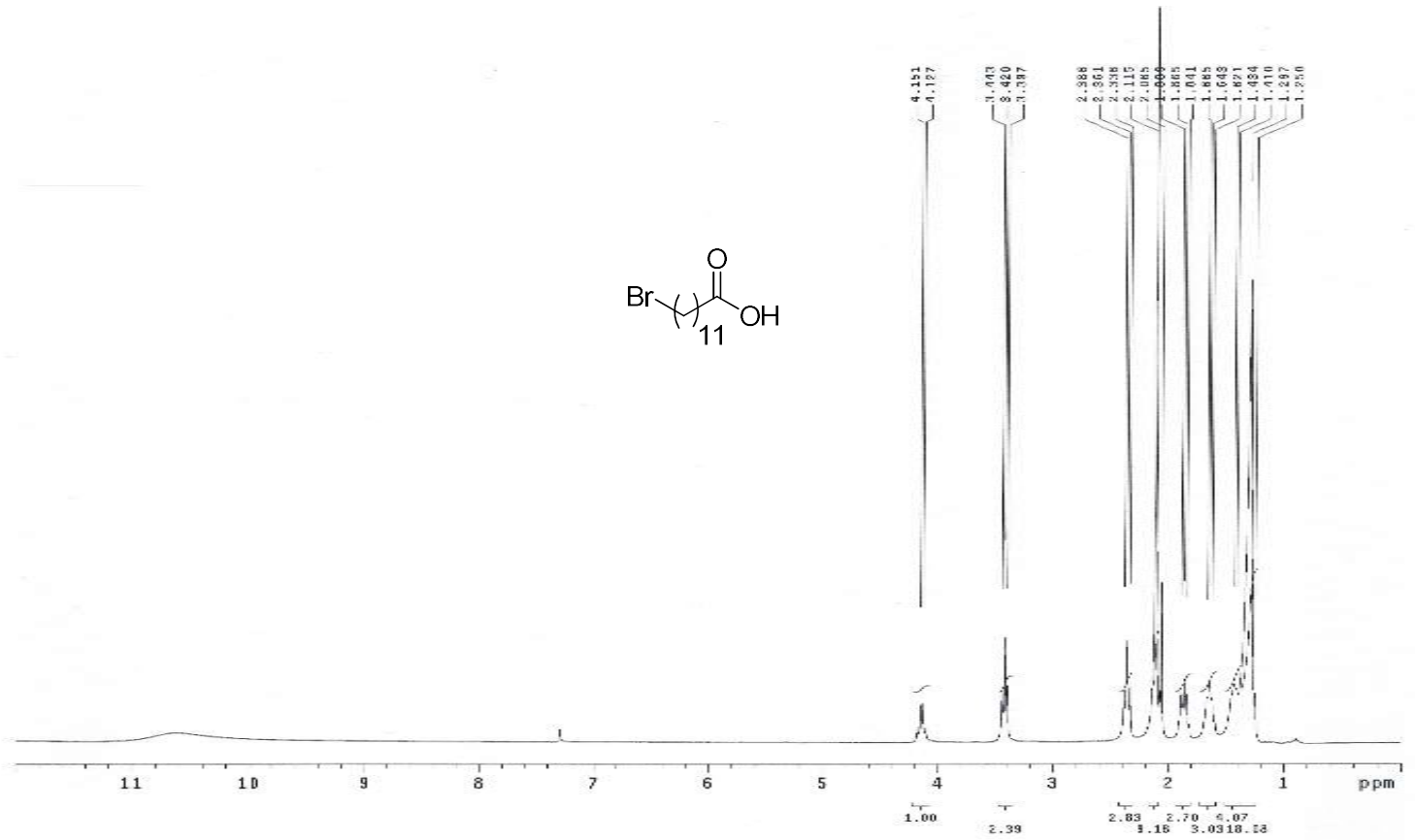


Figure S8: The proton NMR spectra of compound 16.

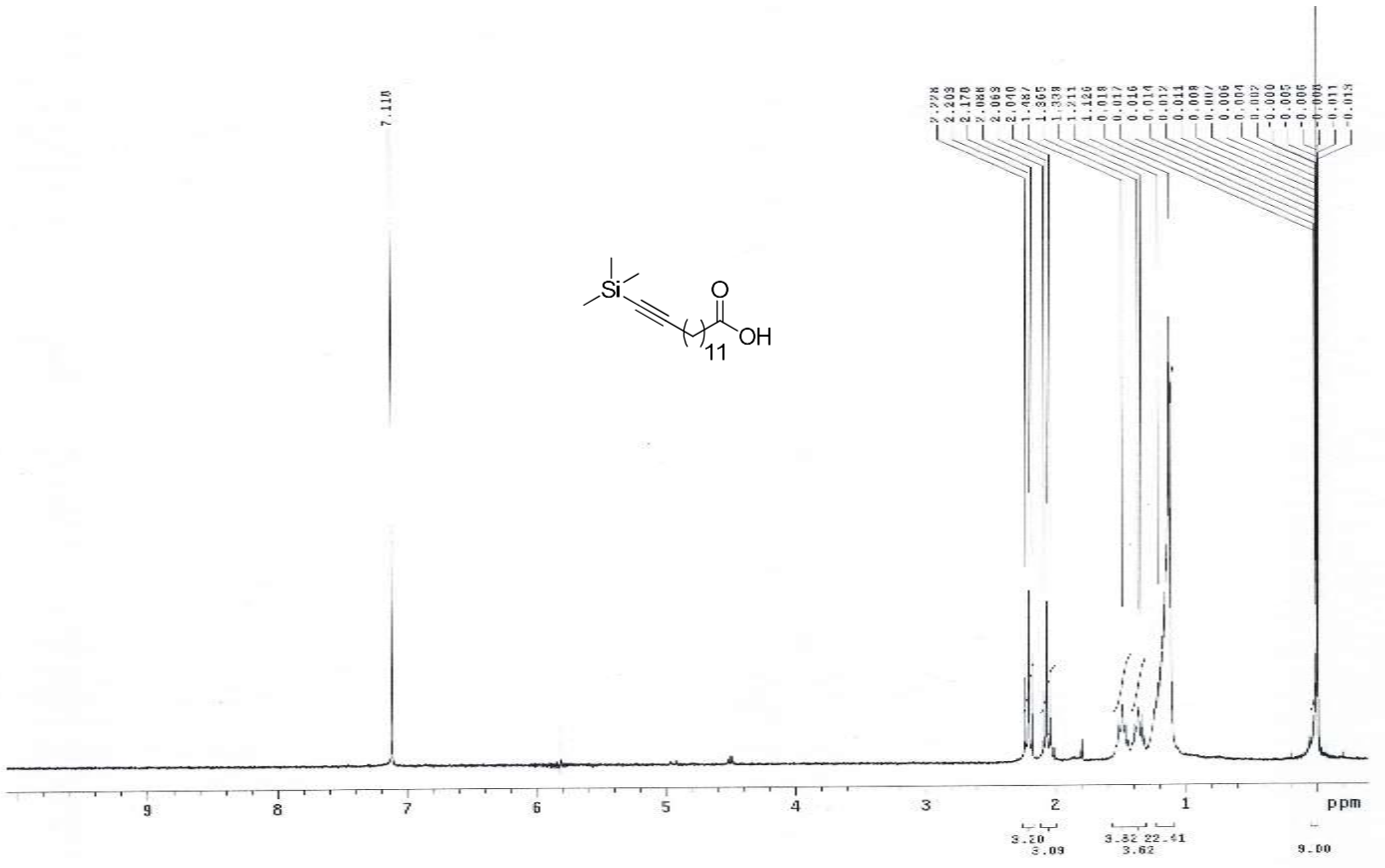


Figure S9: The proton NMR spectra of compound 17.

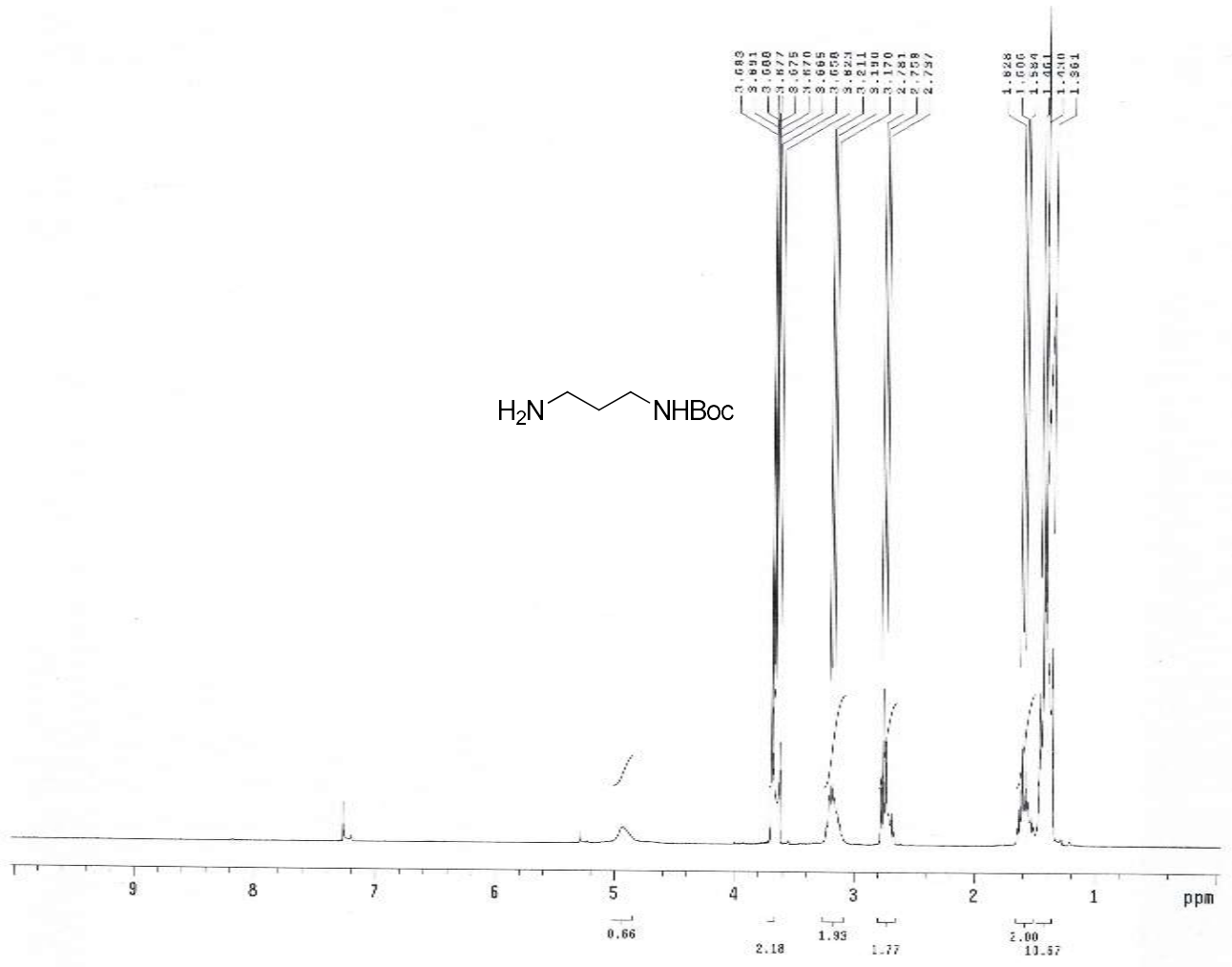


Figure S10: The proton NMR spectra of compound 20.

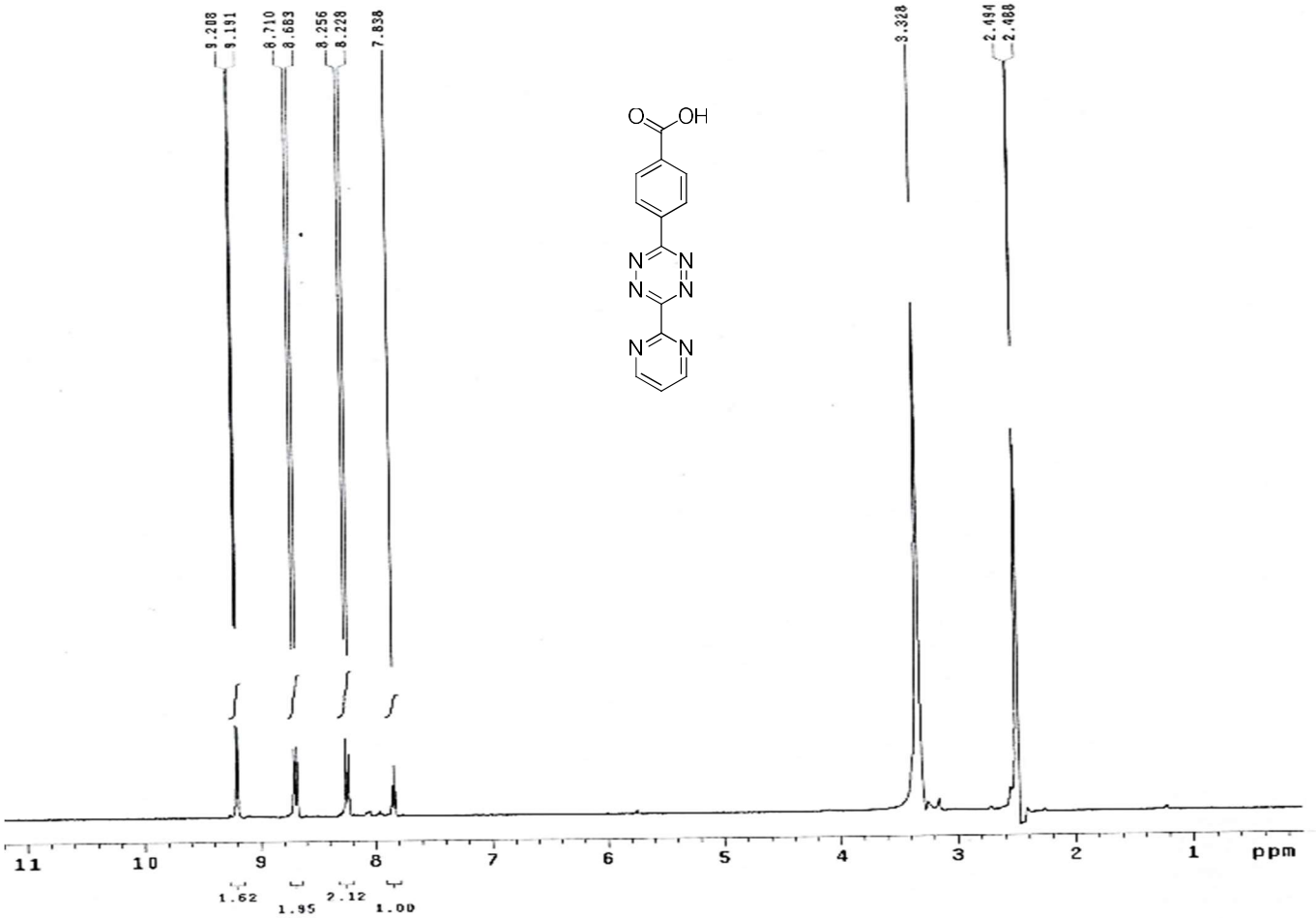


Figure S11: The proton NMR spectra of compound 25.

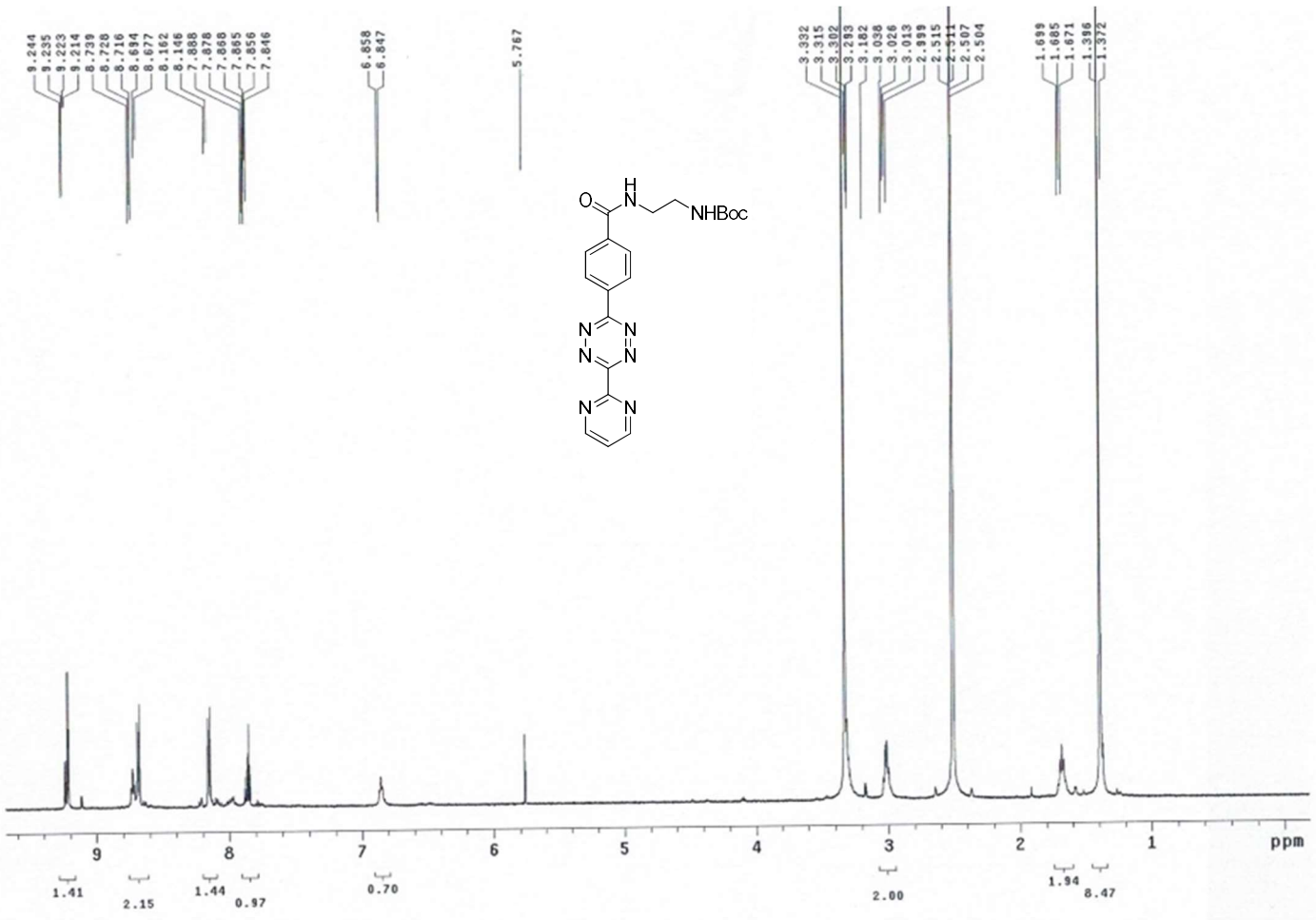


Figure S12: The proton NMR spectra of compound 26.

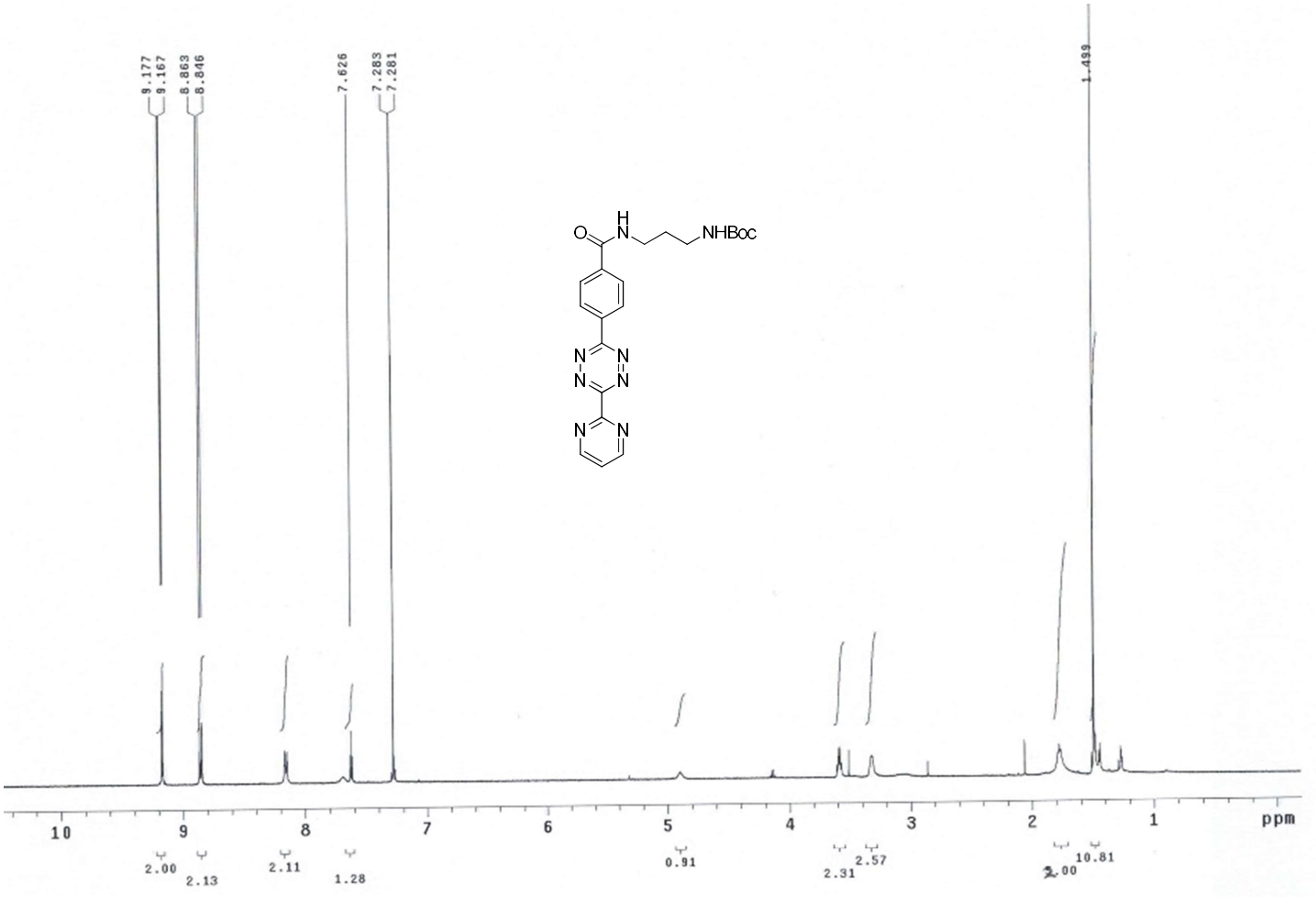


Figure S13: The proton NMR spectra of compound 28.

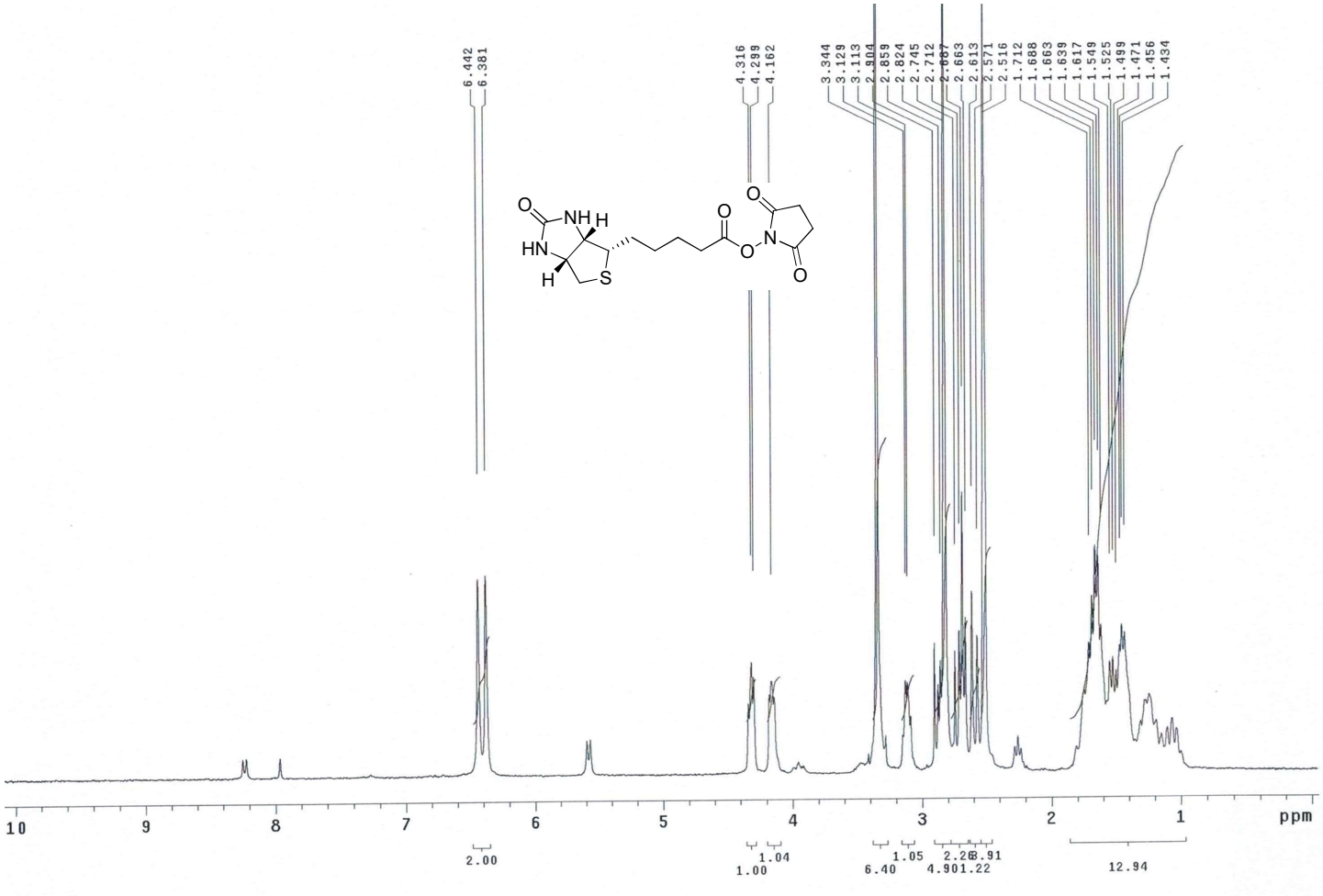


Figure S14: The proton NMR spectra of compound 31.